Universidad Zaragoza

1542

Trabajo Fin de Máster

Estadística 4º de la ESO para matemáticas académicas: una propuesta didáctica.

Statistics. A didactic proposal for the 4th Year of ESO - Academic oriented mathematics.

Autora

Juana Mª Jiménez Ruiz

Directora

Carmen Julve Tiestos

FACULTAD DE EDUCACIÓN

2020

Tabla de contenido

A. Sobre la definición del objeto matemático a enseñar

El objeto matemático que vamos a tratar es estadística de 4º de ESO de matemáticas orientadas a las enseñanzas académicas.

Actualmente la idea general que existe sobre la estadística, debido a la influencia de los medios de comunicación, es la de un conjunto de datos numéricos relacionados que se presentan de forma ordenada y sistemática. Pero la estadística adquiere especial relevancia cuando se aplica en estudios científicos y sociológicos de interés, cuya evolución parece ser incierta dada su variabilidad. La estadística es la herramienta de las matemáticas aplicadas que estudia fenómenos cuyos resultados no son deterministas.

Por ejemplo, la estadística es una ciencia de aplicación práctica:

- **En sondeos electorales** para conocer las preferencias de los electores antes de hacer una votación y así orientar las estrategias de los candidatos;
- **En sanidad** para conocer la evolución de una pandemia y hacer un modelo predictivo;
- **En el mundo empresarial** para tomar decisiones acertadas en función de los costes reales.

Debemos distinguir entre dos tipos de estadísticas:

- **La estadística descriptiva** que representa en gráficos y tablas la información de un conjunto de datos, resumida y analizada mediante cálculos numéricos.
- **La estadística inferencial** que partiendo de datos muestrales efectúa estimaciones, predicciones y generalizaciones sobre un conjunto mayor de datos. Los estudios estadísticos inferenciales permiten establecer relaciones entre las variables y así predecir qué puede ocurrir en el futuro. Esto permite tomar decisiones justas y evitar situaciones de conflicto.

A pesar de la gran variedad de recursos educativos existentes y los avances tecnológicos, la praxis educativa se sigue apoyando fuertemente en el libro de texto lo que le otorga rigidez al proceso de enseñanza-aprendizaje y mucha autoridad al libro. Por esta razón la investigación de los libros de texto es un enclave en la calidad de nuestro sistema educativo.

En mi TFM hago un estudio y análisis comparativo de tres libros de matemáticas de este curso, editados en 2016 por Edelvives, Oxford Educación y Savia SM. En

las figuras 1, 2 y 3 se puede ver como comienzan la estadística con representaciones pictóricas y actividades sobre los contenidos.

***ESO 4 matemáticas académicas* de la editorial Edelvives** hace una presentación de la estadística que ocupa dos páginas donde se combinan tres fotografías que reflejan tres tipos de gráficas estadísticas, cuya representación **muestra elementos de la vida cotidiana** que forman una composición visual análoga a gráficos estadísticos: una tarta circular compuesta por porciones distintas similar a un diagrama de sectores, una pila de lapiceros de distinta longitud que se parece un diagrama de barras o histograma y una distribución de hojas secas que se asemeja a una nube de puntos que sigue un patrón determinado.

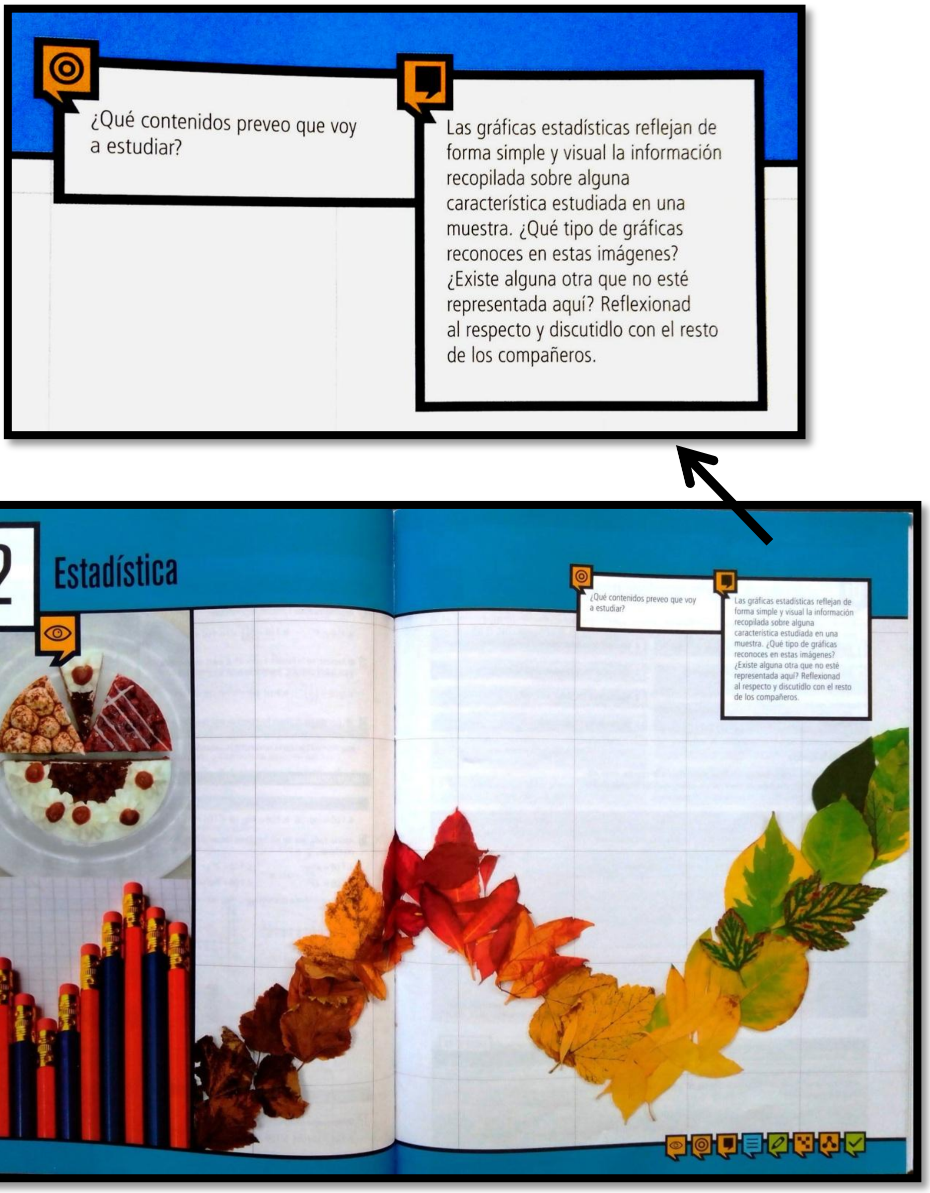
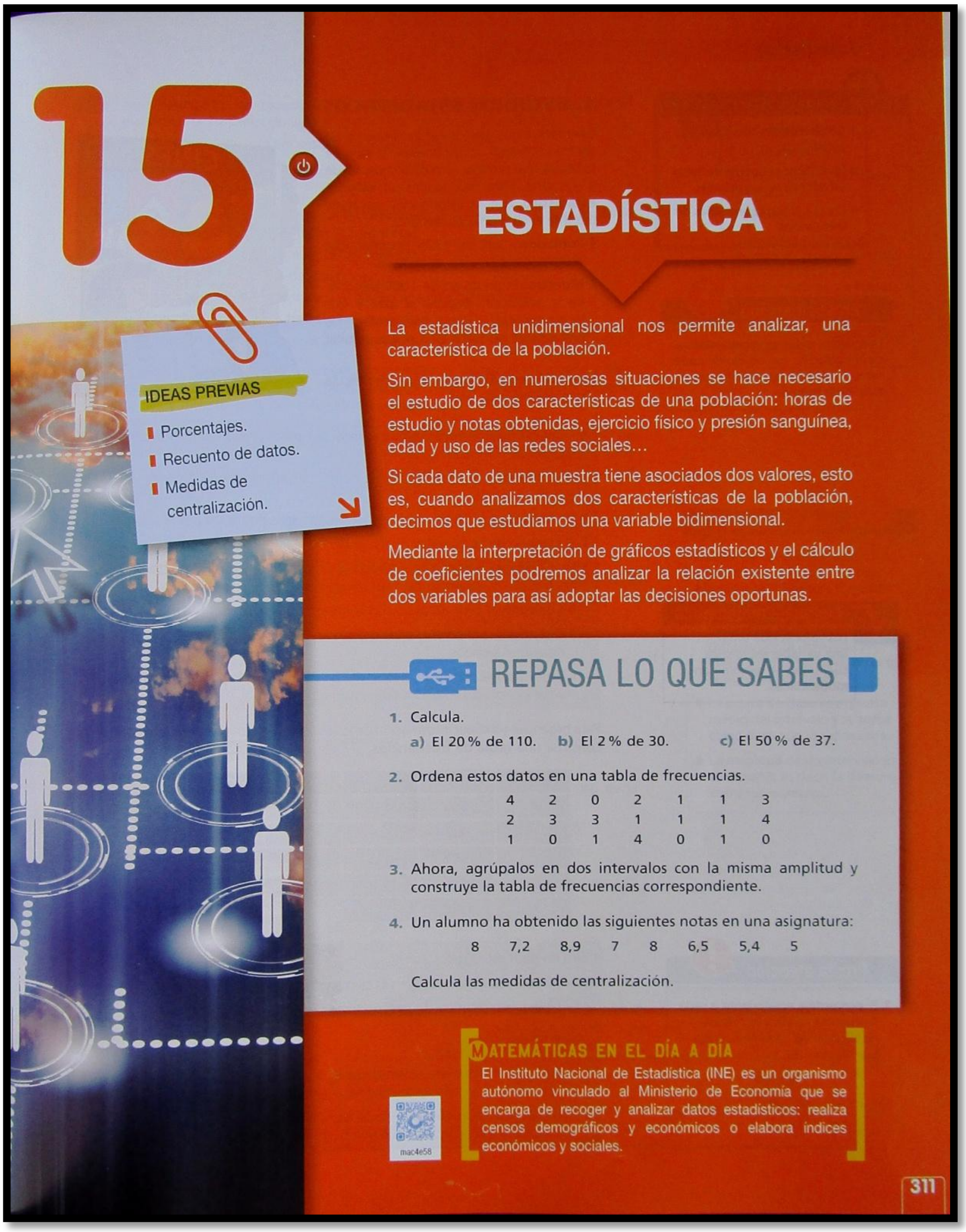


Figura 1. Introducción a la estadística en el libro EDELVIVES.

***Matemáticas orientadas a las enseñanzas académicas 4 ESO* de la editorial Oxford Educación** posee un formato más austero, le dedica una sola página a la portada de la unidad optando por un mayor contenido textual que pictórico. **Presenta una nota de conceptos previos** junto con el enunciado de un problema para repasar el conocimiento adquirido en cursos anteriores. Justifica su aprendizaje como una habilidad que es necesario adquirir para analizar la información que recibimos en la vida cotidiana a través de los distintos medios. Para resaltar su relevancia **nombra al Instituto Nacional de Estadística (INE) vinculado al Ministerio de Economía.**



15

ESTADÍSTICA

IDEAS PREVIAS

- Porcentajes.
- Recuento de datos.
- Medidas de centralización.

La estadística unidimensional nos permite analizar, una característica de la población.

Sin embargo, en numerosas situaciones se hace necesario el estudio de dos características de una población: horas de estudio y notas obtenidas, ejercicio físico y presión sanguínea, edad y uso de las redes sociales…

Si cada dato de una muestra tiene asociados dos valores, esto es, cuando analizamos dos características de la población, decimos que estudiamos una variable bidimensional.

Mediante la interpretación de gráficos estadísticos y el cálculo de coeficientes podremos analizar la relación existente entre dos variables para así adoptar las decisiones oportunas.

REPASA LO QUE SABES

1. Calcula.
   a) El 20 % de 110. b) El 2 % de 30. c) El 50 % de 37.
2. Ordena estos datos en una tabla de frecuencias.

4	2	0	2	1	1	3
2	3	3	1	1	1	4
1	0	1	4	0	1	0

3. Ahora, agrúpalos en dos intervalos con la misma amplitud y construye la tabla de frecuencias correspondiente.
4. Un alumno ha obtenido las siguientes notas en una asignatura:
   8 7,2 8,9 7 8 6,5 5,4 5
   Calcula las medidas de centralización.

MATEMÁTICAS EN EL DÍA A DÍA

El Instituto Nacional de Estadística (INE) es un organismo autónomo vinculado al Ministerio de Economía que se encarga de recoger y analizar datos estadísticos: realiza censos demográficos y económicos o elabora índices económicos y sociales.

mac4e58

311

Figura 2. Introducción a la estadística en el libro OXFORD EDUCACIÓN.

***Matemáticas orientadas a las enseñanzas académicas 4 ESO* de la editorial Savia SM** le dedica dos páginas en las que se observa una imagen de varias torres de monedas apiladas a distinta altura. Estas "torres" reflejan simbólicamente un gráfico estadístico, el histograma. **La imagen tiene doble connotación porque de un modo simbólico relaciona la estadística con la actividad económica.** En la página de la derecha incluye un texto que relata cómo surgen los histogramas en la historia, y añade una serie de cuestiones para reflexionar sobre este texto y la relevancia de la estadística en la actualidad.

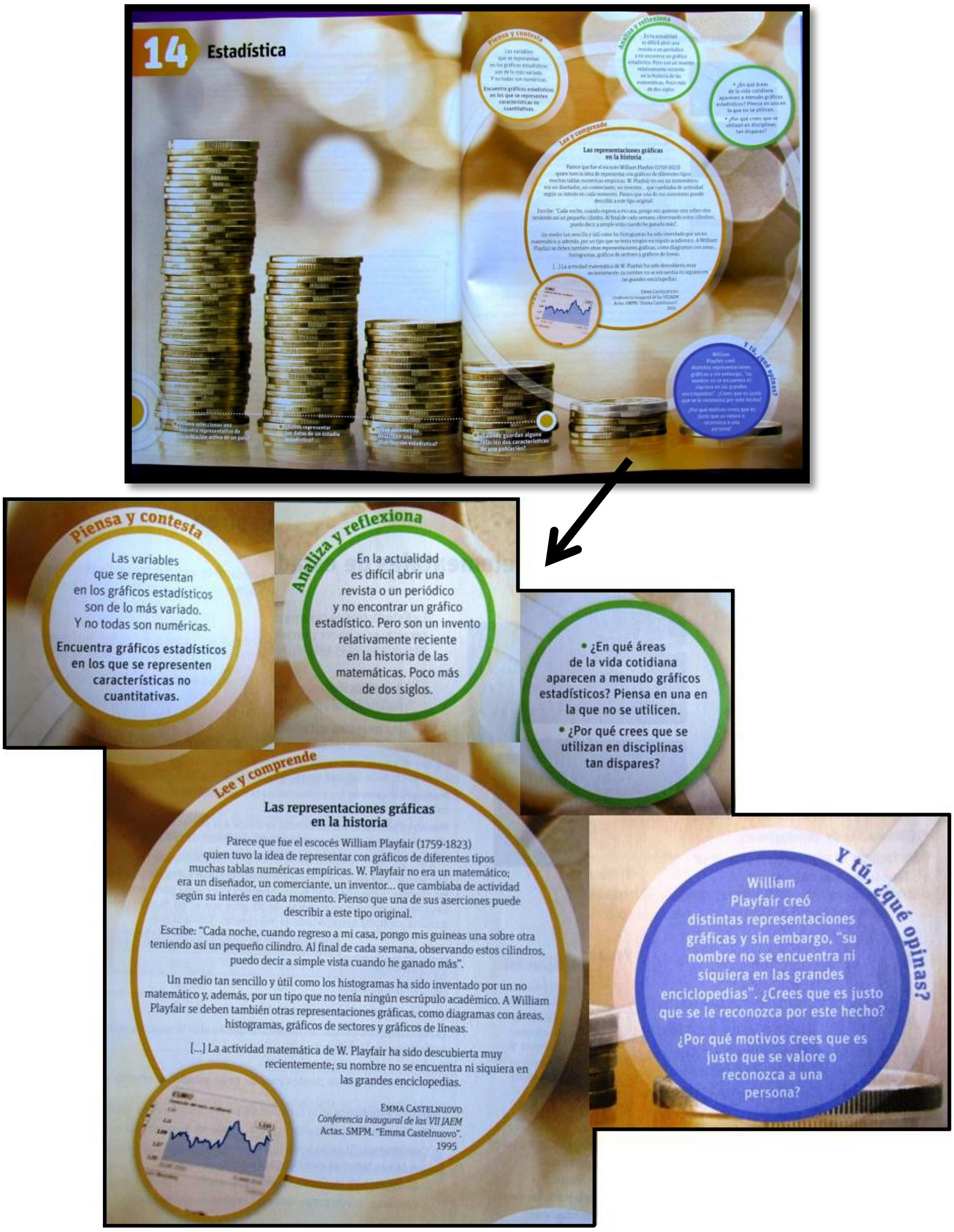


Figura 3. Introducción a la estadística en el libro SAVIA SM.

Si se hace una comparación de las tres introducciones, se observan importantes diferencias en cuanto a formato, contenido y contexto.

De los tres libros de texto considero más apropiada la introducción que ofrece el libro de Savia SM por estas razones:

- El formato visual es más rico y representativo.
- Presenta un contexto histórico que justifica la utilidad y razón de ser de la estadística en la sociedad.
- Pone de manifiesto la estadística como una materia interdisciplinar.
- Las actividades didácticas que presenta invitan al alumno a:
  - investigar sobre los distintos conceptos estadísticos;
  - reflexionar sobre la influencia de la estadística en la vida cotidiana;
  - plantearse cuestiones relacionadas con los aspectos funcionales de la estadística;
  - considerar la estadística como una materia interdisciplinar.
- Además, entre las actividades incluye una que desarrolla del espíritu crítico de los alumnos y que consiste en debatir la siguiente pregunta: "¿Por qué motivos crees que es justo que se valore o reconozca a una persona?"

1. Campo de problemas, técnicas y tecnologías asociadas al objeto matemático

La actividad matemática puede identificarse en su mayoría como una actividad de modelización. **La tabla 1 relaciona los componentes de la actividad matemática: los campos de problemas, técnicas y tecnologías que justifican las técnicas, asociadas a la enseñanza de la estadística y que considero en mi propuesta didáctica.** Suponen la base del desarrollo de mi TFM, además del eje central en torno al cual gira el diseño de mis problemas y actividades con el objetivo de mejorar dicha enseñanza en el aula.

Tabla 1

Campos de problemas, técnicas y tecnologías

Campos de problemas	**Técnicas (praxis)**	**Tecnologías (logos)**
Muestreo. Tablas de datos y de frecuencias. Estudio y extracción de inferencias.	Ordenación de datos y construcción de la tabla con las columnas típicas: N, x_i, f_i, F_i, h_i , H_i y $\sum$. Tablas de frecuencias.	Definición de población y muestra, variable estadística, frecuencias absoluta y relativa (acumuladas o no).
Organización visual de los datos estadísticos.	Elaboración de los diferentes tipos de gráficos.	Definición de diagrama de barras, polígono de frecuencias (acumuladas o no), diagrama de sectores, diagrama lineal e histograma.
Medidas de centralización y de posición.	A partir de las tablas de datos y tablas de frecuencias y aplicando la correspondiente definición de moda, media,… calculamos las diferentes medidas.	Definición de moda, media, mediana y cuartiles.
Medidas de dispersión.	A partir de las tablas de datos y tablas de frecuencias y aplicando la correspondiente definición de recorrido, varianza,... calculamos las diferentes medidas.	Definición de recorrido o rango, recorrido intercuartílico, varianza, desviación típica, coeficiente de variación y distribución normal.
Distribuciones bidimensionales.	A partir de las tablas de frecuencias y de contingencia, diagrama de dispersión o nube de puntos, calculamos los diferentes parámetros como la covarianza o el coeficiente de correlación.	Definición de centro de gravedad, covarianza, coeficiente de correlación y recta de regresión lineal.

B. Sobre el estado de la enseñanza – aprendizaje de la estadística

A nivel estatal las directrices curriculares actuales se enmarcan dentro de la Ley Orgánica 2/2006, de 3 de mayo, de Educación (LOE) modificada por la Ley Orgánica 8/2013, de 9 de diciembre, para la Mejora de la Calidad Educativa (LOMCE), que establece en su Título Preliminar, los principios en los que se inspira el sistema educativo español.

A nivel autonómico el currículo oficial de matemáticas se apoya en la orden ECD/489/2016, de 26 de mayo, por la que se aprueba el currículo de la Educación Secundaria Obligatoria y se autoriza su aplicación en los centros docentes de la Comunidad Autónoma de Aragón.

En estas directrices curriculares los contenidos del área de matemáticas se organizan en cinco bloques, de los cuales el primero es transversal a los demás: 1) Procesos, métodos y actitudes en matemáticas; 2) Números y álgebra; 3) Geometría; 4) Funciones; 5) Probabilidad y estadística.

A continuación especifico los contenidos del bloque transversal y de estadística en los cuales me voy centrar para desarrollar mi trabajo.

BLOQUE 1: Procesos, métodos y actitudes en matemáticas.

Contenidos:

- Planificación del proceso de resolución de problemas.
- Estrategias y procedimientos puestos en práctica: uso del lenguaje apropiado (gráfico, numérico, algebraico, etc.), reformulación del problema, resolver subproblemas, recuento exhaustivo, empezar por casos particulares sencillos, buscar regularidades y leyes, etc.
- Reflexión sobre los resultados: revisión de las operaciones utilizadas, asignación de unidades a los resultados, comprobación e interpretación de las soluciones en el contexto de la situación, búsqueda de otras formas de resolución, etc.
- Planteamiento de investigaciones matemáticas escolares en contextos numéricos, geométricos, funcionales, estadísticos y probabilísticos.
- Práctica de los procesos de matematización y modelización, en contextos de la realidad y en contextos matemáticos.

- Confianza en las propias capacidades para desarrollar actitudes adecuadas y afrontar las dificultades propias del trabajo científico.
- Utilización de medios tecnológicos en el proceso de aprendizaje para:
  a) la recogida ordenada y la organización de datos;
  b) la elaboración y creación de representaciones gráficas de datos numéricos, funcionales o estadísticos;
  c) facilitar la comprensión de propiedades geométricas o funcionales y la realización de cálculos de tipo numérico, algebraico o estadístico;
  d) el diseño de simulaciones y la elaboración de predicciones sobre situaciones matemáticas diversas;
  e) la elaboración de informes y documentos sobre los procesos llevados a cabo y los resultados y conclusiones obtenidos;
  f) comunicar y compartir, en entornos apropiados, la información y las ideas matemáticas.

BLOQUE 5: Estadística.

Contenidos:

- Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar y la estadística.
- Identificación de las fases y tareas de un estudio estadístico.
- Gráficas estadísticas: distintos tipos de gráficas. Análisis crítico de tablas y gráficas estadísticas en los medios de comunicación. Detección de falacias.
- Medidas de centralización y dispersión: interpretación, análisis y utilización.
- Comparación de distribuciones mediante el uso conjunto de medidas de posición y dispersión.
- Construcción e interpretación de diagramas de dispersión. Introducción a la correlación.

Criterios de evaluación, competencias clave y estándares de aprendizaje.

Crit.MAAC.5.3. Utilizar el lenguaje adecuado para la descripción de datos y analizar e interpretar datos estadísticos que aparecen en los medios de comunicación. CCL-CMCT

  - Est.MAAC.5.3.1. Utiliza un vocabulario adecuado para describir, cuantificar y analizar situaciones relacionadas con el azar.

Crit.MAAC.5.4. Elaborar e interpretar tablas y gráficos estadísticos, así como los parámetros estadísticos más usuales, en distribuciones unidimensionales y bidimensionales, utilizando los medios más adecuados (lápiz y papel, calculadora u ordenador), y valorando cualitativamente la representatividad de las muestras utilizadas. CMCT-CD-CAA

- Est.MAAC.5.4.1. Interpreta críticamente datos de tablas y gráficos estadísticos.
- Est.MAAC.5.4.2. Representa datos mediante tablas y gráficos estadísticos utilizando los medios tecnológicos más adecuados.
- Est.MAAC.5.4.3. Calcula e interpreta los parámetros estadísticos de una distribución de datos utilizando los medios más adecuados (lápiz y papel, calculadora u ordenador).
- Est.MAAC.5.4.4. Selecciona una muestra aleatoria y valora la representatividad de la misma en muestras muy pequeñas.
- Est.MAAC.5.4.5. Representa diagramas de dispersión e interpreta la relación existente entre las variables.

1. Los libros de texto

He realizado un análisis de los tres libros de texto de 4º de ESO del año 2016. Estos libros de texto han sido prestados por el CPRI-ES La Salle – Franciscanas Gran Vía para la realización de mi TFM. **La tabla 2 muestra en qué orden los libros de texto escolares introducen los contenidos, se ve como siguen secuencialmente el currículo oficial dejando la estadística para el final de curso.** La editorial Edelvives la antepone a los contenidos de probabilidad, mientras que las otras editoriales la sitúan como la última unidad.

Tabla 2

Los bloques de contenido y las unidades didácticas en los libros de texto

Bloques de contenido	Libro 1 (14 uds.) **Edelvives**	Libro 2 (15 uds.) **Oxford Educación**	Libro 3 (14 uds.) **Savia SM**
Números y álgebra.	Unidades 1 - 4	Unidades 1 - 5	Unidades 1 - 4
Geometría.	Unidades 5 - 9	Unidades 6 - 8	Unidades 5 - 7
Funciones.	Unidades 10 - 11	Unidades 9 - 12	Unidades 8 - 11
Probabilidad y estadística.	**Estadística - 12**	Combinatoria - 13	Combinatoria - 12
	Combinatoria - 13	Probabilidad - 14	Probabilidad - 13
	Probabilidad - 14	**Estadística - 15**	**Estadística - 14**

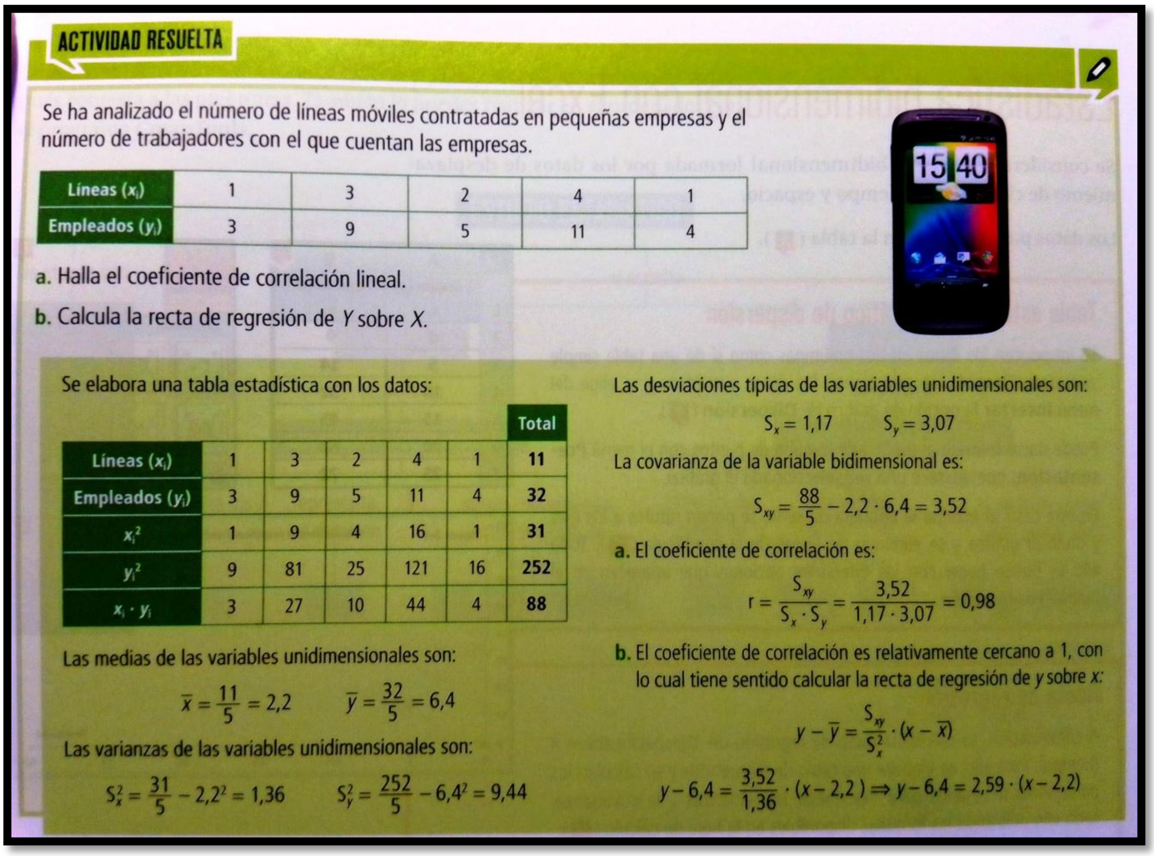
Mi análisis de los libros de texto se basa en el estudio de tres aspectos:

- los contenidos matemáticos que se presentan;
- el contenido didáctico (ejemplos, actividades, recursos TIC, etc.);
- el diseño característico de cada editorial.

Primero, he realizado un análisis cualitativo y cuantitativo de cada uno de estos tres libros, y a continuación, un análisis comparativo.

<u>Libro 1 Edelvives:</u> ***ESO 4 matemáticas académicas, volumen 3***

La unidad de estadística está subdividida en 7 lecciones que introducen, de forma jerarquizada, los conceptos estadísticos en 16 páginas. **Tiene un formato en el que predominan las definiciones textuales y visuales de los conceptos básicos de estadística frente a los símbolos y fórmulas matemáticas que los representan. Recopila un total de 45 actividades que se presentan paralelamente a la introducción teórica para practicar, algunas de ellas ya resueltas (ver tabla 3).** En las 4 primeras lecciones la mayoría de las actividades están contextualizadas, pero en las 3 últimas, la parte correspondiente a la correlación entre dos variables, casi todos los ejercicios son de aplicación directa de las técnicas correspondientes, sobre interpretación de diagramas de dispersión y valores de coeficiente de correlación, o cuestiones para afianzar el grado de asimilación de los conceptos estadísticos.



ACTIVIDAD RESUELTA

Se ha analizado el número de líneas móviles contratadas en pequeñas empresas y el número de trabajadores con el que cuentan las empresas.

Líneas (x_i)	1	3	2	4	1
Empleados (y_i)	3	9	5	11	4

a. Halla el coeficiente de correlación lineal.

b. Calcula la recta de regresión de Y sobre X.

Se elabora una tabla estadística con los datos:

						Total
Líneas (x_i)	1	3	2	4	1	11
Empleados (y_i)	3	9	5	11	4	32
x_i^2	1	9	4	16	1	31
y_i^2	9	81	25	121	16	252
$x_i \cdot y_i$	3	27	10	44	4	88

Las medias de las variables unidimensionales son:

$$\bar{x} = \frac{11}{5} = 2,2 \qquad \bar{y} = \frac{32}{5} = 6,4$$

Las varianzas de las variables unidimensionales son:

$$S_x^2 = \frac{31}{5} - 2,2^2 = 1,36 \qquad S_y^2 = \frac{252}{5} - 6,4^2 = 9,44$$

Las desviaciones típicas de las variables unidimensionales son:

$$S_x = 1,17 \qquad S_y = 3,07$$

La covarianza de la variable bidimensional es:

$$S_{xy} = \frac{88}{5} - 2,2 \cdot 6,4 = 3,52$$

a. El coeficiente de correlación es:

$$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{3,52}{1,17 \cdot 3,07} = 0,98$$

b. El coeficiente de correlación es relativamente cercano a 1, con lo cual tiene sentido calcular la recta de regresión de y sobre x:

$$y - \bar{y} = \frac{S_{xy}}{S_x^2} \cdot (x - \bar{x})$$

$$y - 6,4 = \frac{3,52}{1,36} \cdot (x - 2,2) \Rightarrow y - 6,4 = 2,59 \cdot (x - 2,2)$$

Figura 4. Actividad resuelta del libro EDELVIVES.

Un ejemplo de actividad resuelta es la figura 4. En el desarrollo simplemente se da el valor de los parámetros estadísticos pero se debería indicar qué sentido tienen respecto al problema. Por ejemplo, se debería decir que las empresas analizadas tienen una media de 6,4 empleados. No se explica que el coeficiente de correlación R=0,98 indica una fuerte relación entre el número de empleados y el número de líneas contratadas, de este modo podemos calcular a través de la recta de regresión el número de líneas contratadas según el número de empleados que tenga el empresa.

Tabla 3

Actividades de la unidad de estadística en el libro EDELVIVES

Actividades a lo largo de la unidad	Nº de actividades	Contextualizado	Total
Lección 1. (1 página) **Etapas de un estudio estadístico.**	1 resuelta	1	1
Lección 2. (3 páginas) **Estadística unidimensional.**	3	2	4
**http://www.ine.es/explica_pasos_primera_encuesta.htm*	1 enlace*	¿?	
Lección 3. (4 páginas) **Estadística bidimensional.**	2 resueltas	1	11
	9	8	
Lección 4. (2 páginas) **Dependencia aleatoria y funcional.**	9	7	9
Lección 5. (2 páginas) **Correlación entre dos variables.**	1 resuelta	1	6
	5	2	
Lección 6. (2 páginas) **Coeficiente de correlación lineal. Interpretación.**	9	3	9
Lección 7. (2 páginas) **Regresión lineal. Rectas de regresión.**	1 resuelta	1	5
	4	1	
			TOTAL = 45

Por un lado, los conceptos teóricos más relevantes están resaltados en color amarillo, azul y marrón, por otro lado, las actividades en verde. Llama la atención el grado de intensidad que alcanzan estos colores en algunas secciones, diferenciando mucho el contenido teórico del práctico. **Otra característica de este libro es el formato tabular que presenta, donde combina tablas con encabezados en vertical y horizontal**, lo que implica que muchos cálculos de sumas totales se hagan en uno y en otro sentido, cuando tradicionalmente en las cabeceras verticales de las tablas estadísticas se nombran simbólicamente los parámetros estadísticos (variables, marcas de clase, frecuencias, porcentajes, etc.).

Otro ejemplo de problema descontextualizado se observa en la figura 5. De nuevo al leer el enunciado del problema no se requiere hacer inferencias concretas, solo se pide hacer el cálculo de la distribución marginal.

7 ⊕ Para comprobar el ahorro que se puede hacer en la compra diaria, se ha registrado el precio de una barra de pan y de un litro de leche, en céntimos de euro, en diferentes supermercados de una ciudad:

(45 , 79) (40 , 77) (40 , 78) (45 , 77) (50 , 79)
(35 , 77) (55 , 75) (45 , 76) (45 , 79) (40 , 75)
(50 , 75) (45 , 76) (45 , 79) (35 , 78) (55 , 79)

a. Recoge los valores en una tabla de doble entrada.

b. Representa el diagrama de dispersión de los puntos.

c. Halla la distribución marginal de las variables X e Y.

Figura 5. Problema del libro EDELVIVES.

Al final de las lecciones dedica una página al uso de herramientas tecnológicas donde explica detalladamente cómo se tratan los datos de la estadística bidimensional con la hoja de cálculo: tabla estadística, gráfico de dispersión y recta de regresión. A continuación ilustra un mapa conceptual para completar en papel o mediante el programa **CmapTools**, una serie de cuestiones teóricas y una propuesta de presentación digital que fomenta el uso del programas tales como **PowerPoint, Glogster,…**

Para terminar, añade una serie de actividades de repaso final de la unidad y de autoevaluación. Estas actividades suman un total de 17 (ver tabla 4).

Tabla 4

Actividades finales en el libro EDELVIVES

Actividades de repaso final	Nº de actividades	Contextualizado	Total
Etapas de un estudio estadístico.	1	1	1
Estadística unidimensional.	4	4	4
Estadística bidimensional.	2	2	2
Dependencia aleatoria y funcional.	1	1	1
Correlación entre dos variables.	1	1	1
Coeficiente de correlación lineal. Interpretación.	2	2	2
Regresión lineal. Rectas de regresión.	1	1	1
EVALUACIÓN.	5	0	5
			TOTAL = 17

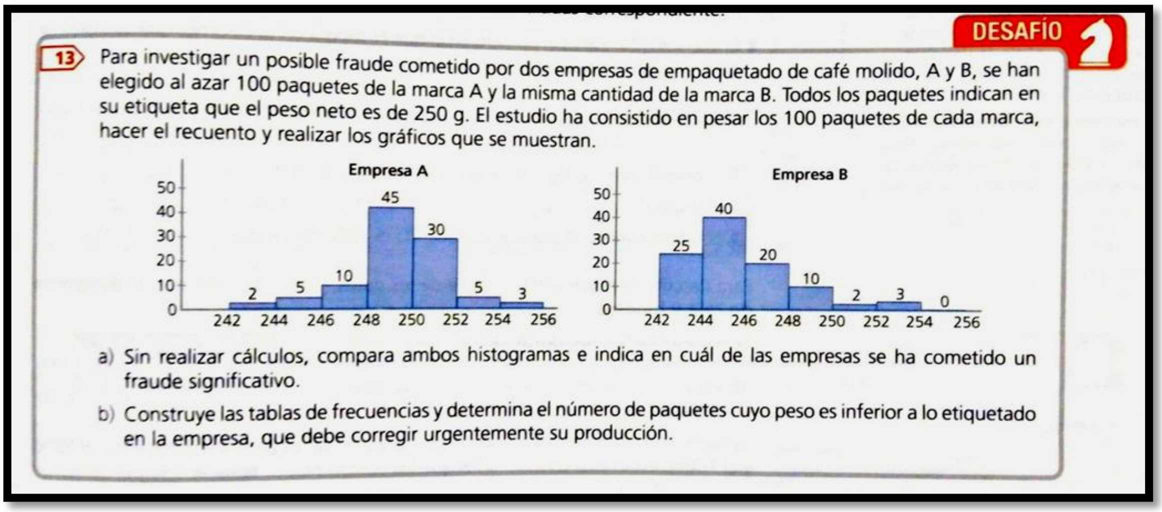
<u>Libro 2 Oxford Educación:</u> *Matemáticas orientadas a las enseñanzas académicas 4 ESO, volumen 3*

En este libro la unidad está subdividida en 6 lecciones donde se introducen los conceptos estadísticos progresivamente en 12 páginas. Cada lección se extiende a 2 páginas, de las cuales 1 página se dedica exclusivamente al enunciado de actividades para resolver.

Presenta grandes diferencias en comparación a Edelvives: involucra más conceptos estadísticos, más simbología matemática, fórmulas desarrolladas, el formato tabular estadístico a modo tradicional y los gráficos son más reducidos en tamaño que en el anterior. Estas diferencias que son muy significativas entre los libros de texto, las expongo explícitamente más adelante cuando hago un análisis comparativo más exhaustivo y en profundidad de los tres libros.

Oxford Educación está editado en colores claros y con muchos menos contrastes, aunque las ideas principales se resaltan igualmente. El libro combina un predominante fondo blanco con cuadros en tonos grises, y notas con el texto del título en blanco resaltado sobre fondo azul claro.

En el problema de la figura 6 tenemos un enunciado que hace referencia expresa a un contexto relacionado con la productividad de una empresa. En este caso no se pide sin más el cálculo de una medida o parámetro, sino que hay responder de manera razonada qué empresa ha cometido mayor fraude y calcular el número de paquetes cuyo peso es inferior a lo etiquetado.


DESAFÍO

13 Para investigar un posible fraude cometido por dos empresas de empaquetado de café molido, A y B, se han elegido al azar 100 paquetes de la marca A y la misma cantidad de la marca B. Todos los paquetes indican en su etiqueta que el peso neto es de 250 g. El estudio ha consistido en pesar los 100 paquetes de cada marca, hacer el recuento y realizar los gráficos que se muestran.

a) Sin realizar cálculos, compara ambos histogramas e indica en cuál de las empresas se ha cometido un fraude significativo.

b) Construye las tablas de frecuencias y determina el número de paquetes cuyo peso es inferior a lo etiquetado en la empresa, que debe corregir urgentemente su producción.

Figura 6. Problema "desafío" en el libro OXFORD EDUCACIÓN.

La mayoría de las actividades se encuentran en un contexto y se específica en una escala que va del 1 al 3 el nivel de dificultad que poseen a la hora de resolver: nivel 1 bajo, nivel 2 medio y nivel 3 alto. Además de un "desafío" para atender la diversidad en el aula. Son un total de 41 actividades (ver tabla 5).

Tabla 5

Actividades de la unidad de estadística en el libro OXFORD EDUCACIÓN

Actividades a lo largo de la unidad	**Nº de actividades (nivel)**	**Contextualizado**	**Total**
Lección 1. (2 páginas) **Estudios estadísticos.**	4 (1)	3	8
	3 (2)	3	
	1 desafío	1	
Lección 2. (2 páginas) **Gráficos estadísticos.**	1 (1)	1	5
	2 (2)	2	
	1 (3)	1	
	1 desafío	1	
Lección 3. (2 páginas) **Medidas de centralización y de posición.**	5 (1)	3	8
	2 (2)	2	
	1 desafío	0	
Lección 4. (2 páginas) **Medidas de dispersión.**	3 (1)	2	6
	1 (2)	1	
	1 (3)	1	
	1 desafío	1	
Lección 5. (2 páginas) **Variables estadísticas bidimensionales. Diagramas de dispersión.**	3 resueltas	1	9
	4 (1)	4	
	1 (2)	1	
	1 desafío	0	
Lección 6. (2 páginas) **Covarianza. Correlación lineal.**	1 (1)	1	5
	3 (2)	1	
	1 desafío	1	
			TOTAL = 41

En la figura 7 tenemos un ejemplo de problema contextualizado donde solo se requiere hacer el cálculo de los parámetros estadísticos, pero que sirve para poner en práctica las técnicas aprendidas y perfeccionarlas.

47 El número de hijos que tienen las familias de una comunidad de vecinos viene representado en la tabla.

Número de hijos	0	1	2	3	4	8
Número de familias	4	5	7	6	2	1

a) Calcula los cuartiles.

b) Dibuja el diagrama de cajas y bigotes.

Figura 7. Problema en el libro OXFORD EDUCACIÓN.

Aparte de las actividades anteriores según aparecen los objetos matemáticos de la estadística, **al final de la unidad se ofrecen un total de 31 actividades para que el estudiante ponga en práctica todo lo aprendido en la lección (ver tabla 6).**

A continuación, la unidad tiene un apartado llamado **"Matemáticas vivas"** que se extiende a 2 páginas y donde se trabaja la aplicación de las matemáticas a partir de situaciones de la vida cotidiana. Concretamente en esta unidad **se estudian las encuestas con intención de voto: los sondeos realizados para unas elecciones generales**, razonar los datos de esta encuesta, reflexionar sobre posibles alianzas entre los diferentes partidos así como reflexionar sobre la distribución de votos por tramos de edad. Se promueve el trabajo cooperativo y la utilización de las TIC en este proyecto. Se indica la utilización del programa **GeoGebra** para realizar los gráficos estadísticos pero no se explica cómo hacerlos, a diferencia de Edelvives donde sí se indicaban los pasos a seguir para usar la hoja de cálculo Excel.

Por último, el cierre de la unidad se hace con la sección **"Avanza"** para ampliar los conceptos estadísticos: se introducen contenidos de mayor dificultad y otro tipo de gráficos estadísticos que aparecen en la vida cotidiana**. Se comenta como es frecuente utilizar diagramas de tallo y hojas para representar los horarios de servicio en**

estaciones de tren y servicios. Estos diagramas no suelen aparecer en los libros de texto y no aparecen en el currículo oficial. En un principio pensé considerarlos en las técnicas de mi propuesta didáctica pero como suponen una ampliación del currículo, lo desestimé. Sí he de señalar que este tipo de extras suponen una atención a la diversidad al igual que los ejercicios de desafío, y son distintivos de este libro. El uso de las TIC es un medio de atención a la diversidad igualmente pero está más extendido su uso. La inserción de las TIC en los libros de texto es algo habitual, por no decir obligatorio, debido a ello, lo único que los puede diferenciar es la preferencia por apoyarse más en un programa u otro, y la importancia que le dan a su aplicación en el aula.

Al final del volumen, para desarrollar la comprensión lectora desde las matemáticas así como la resolución de problemas, se analizan noticias y artículos en una sección que engloba los contenidos de los bloques de funciones, estadística y probabilidad.

Tabla 6

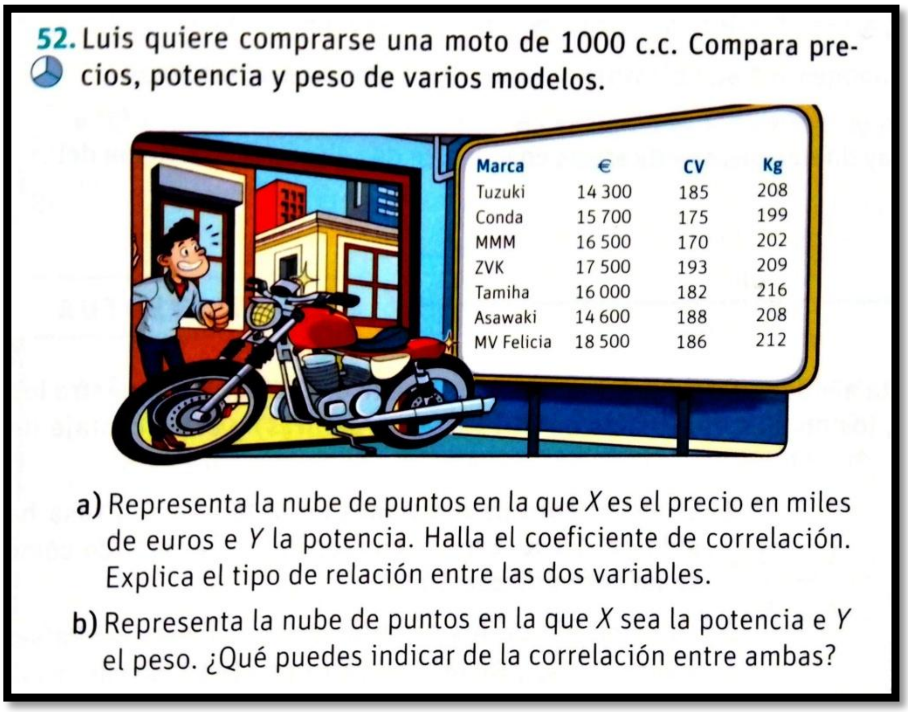
Actividades finales en el libro OXFORD EDUCACIÓN

Actividades finales	Nº de actividades (nivel)	Contextualizado	Total
Estudios y gráficos estadísticos.	6 (1)	5	6
Medidas de centralización y posición.	4 (1)	4	6
	2 (2)	1	
Medidas de dispersión.	5 (1)	4	5
Variables bidimensionales.	3 (1)	2	5
	2 (2)	1	
Covarianza y coeficiente de correlación.	3 (1)	1	9
	4 (2)	4	
	2 (3)	2	
			TOTAL = 31

<u>Libro 3 Savia SM:</u> ***Matemáticas orientadas a las enseñanzas académicas 4 ESO***

La unidad está subdividida en 7 lecciones que se extienden a 13 páginas. A lo largo de las lecciones los contenidos y conceptos más importantes se resaltan en amarillo, las actividades en verde y las tablas en azul. Emplea estos colores en un tono muy claro sin recargar las páginas. En cada lección inserta ejemplos por cada contenido donde aplica y relaciona los conocimientos que se van introduciendo de un modo progresivo. **Por cada lección presenta 2-3 ejemplos resueltos, este libro tiene un marcado carácter ejemplificador que lo distingue de los anteriores. Predominan en cantidad y tamaño las tablas de datos, las tablas de frecuencias y las tablas de contingencia frente a los gráficos.** En los márgenes exteriores de las páginas del libro para cada lección se propone:

- trabajar con distintas estructuras de pensamiento crítico bajo el epígrafe **"Pon en valor",**
- prácticas con actividades interactivas en **smSaviadigital.com,**
- enlaces a actividades interactivas a realizar con **MAT-TIC GeoGebra** en *smSaviadigital.com* para practicar y comprender mejor los conceptos,
- **textos breves** que amplían la información relacionada con la lección,
- indicaciones para hacer cálculos estadísticos con la **calculadora.**



52. Luis quiere comprarse una moto de 1000 c.c. Compara precios, potencia y peso de varios modelos.

Marca	€	CV	Kg
Tuzuki	14 300	185	208
Conda	15 700	175	199
MMM	16 500	170	202
ZVK	17 500	193	209
Tamiha	16 000	182	216
Asawaki	14 600	188	208
MV Felicia	18 500	186	212

a) Representa la nube de puntos en la que X es el precio en miles de euros e Y la potencia. Halla el coeficiente de correlación. Explica el tipo de relación entre las dos variables.

b) Representa la nube de puntos en la que X sea la potencia e Y el peso. ¿Qué puedes indicar de la correlación entre ambas?

Figura 8. Problema en el libro SAVIA SM.

En las figuras 8 y 9 podemos ver cómo son los problemas típicos del Savia SM. En los problemas se involucra un enunciado contextualizado con el cálculo de los parámetros y el desarrollo del sentido estadístico.

Encuentra el error

58. Una imagen vale más que mil tablas.

A veces los parámetros estadísticos pueden jugar malas pasadas. El estadístico Francis J. Anscombe se empeñó en demostrarlo con imágenes.

Observa las dos tablas que inventó.

x	y
4	4,26
5	5,68
6	7,24
7	4,82
8	6,95
9	8,81
10	8,04
11	8,33
12	10,84
13	7,58
14	9,96

x	y
4	3,1
5	4,74
6	6
7	7,26
8	8,14
9	8,87
10	9,14
11	9,26
12	9,13
13	8,74
14	8,1

- La media de la variable y en las dos distribuciones es 7,5.
- En ambas, la variable x toma los mismos valores.
- En las dos el coeficiente de correlación es $r = 0{,}816$.
- La recta de regresión de las dos es: $y = 0{,}5x + 3$.

Y sin embargo, si se afirma que el valor estimado de y para $x = 16$ es 11 se comete un error grave en una de las dos.

¿Por qué? Dibuja las dos nubes de puntos y explica la situación.

Figura 9. Problema del libro SAVIA SM.

A lo largo de la unidad intercala menos actividades con los aspectos teóricos en comparación con Edelvives y Oxford Educación, 21 en total, pero lo compensa con 44 actividades más y otros ejercicios de ampliación al final de la unidad. **Al igual que Oxford Educación, clasifica las actividades por nivel de dificultad del 1 al 3. La mayor parte de los enunciados están contextualizados y hay ejercicios resueltos (ver tablas 7 y 8).**

Tabla 7

Actividades de la unidad de estadística en el libro SAVIA SM

Actividades a lo largo de la unidad	Nº de actividades (nivel)	Contextualizado	Total
Lección 1. (2 páginas) **Conceptos elementales de estadística.**	3 (1)	3	3
Lección 2. (2 páginas) **Gráficos estadísticos.**	1 (1)	1	2
	1 (2)	1	
Lección 3. (2 páginas) **Medidas de centralización.**	1 resuelta	1	4
	3 (1)	3	
Lección 4. (2 páginas) **Medidas de dispersión. Interpretación conjunta de la media y la desviación típica.**	3 (1)	3	4
	1 (2)	1	
Lección 5. (2 páginas) **Distribuciones bidimensionales.**	1 resuelta	1	3
	1 (1)	0	
	1 (2)	1	
Lección 6. (2 páginas) **Covarianza y coeficiente de correlación lineal.**	2 (1)	2	3
	1 (2)	1	
Lección 7. (1 páginas) **Recta de regresión lineal.**	1 (1)	1	2
	1 (2)	1	
			TOTAL = 21

Al final de la unidad se recoge todo su contenido en una página con los aspectos más importantes de la lección. **El resumen "Organiza tus ideas"** presenta la lección subdividida en 5 esquemas independientes que resaltan los contenidos más importantes en este orden: conceptos básicos, gráficos estadísticos, medidas de centralización, medidas de dispersión y distribuciones bidimensionales.

A continuación del resumen de la lección se presentan 2 páginas con 3 **“Actividades clave” resueltas** y resaltadas en fondo azul. Estas actividades son de dos tipos: las más representativas de la unidad o aquellas que muestran procedimientos que permiten relacionar los distintos conceptos estadísticos que se han introducido en la unidad. Son las siguientes:

- *Actividad 1.* Hacer una tabla de frecuencias, representar gráficamente los datos y hallar las medidas de centralización.
- *Actividad 2.* Hacer una tabla de frecuencias de una distribución de datos en intervalos y calcular las medidas de dispersión.
- *Actividad 3.* Hacer la representación gráfica y tabular de una distribución bidimensional. Cuantificar e interpretar la correlación entre las dos variables de la distribución. Hallar la recta de regresión y hacer una estimación del valor de y para un determinado valor de la variable x.

Además, el libro añade 2 páginas más de **“Ponte a prueba” con actividades y problemas de ampliación; tratan temas estadísticos de actualidad extraídos de las pruebas PISA o inspirados en ellos, y una autoevaluación. En las actividades de ampliación se pide una interpretación significativa de los resultados obtenidos.**

En concreto, los 3 problemas inspirados en PISA inducen a la investigación en los siguientes aspectos cívico-sociales:

- *¿Influye el género en los índices de colesterol?*
- *Los niveles de CO2.*
- *Abandono escolar y gasto por estudiante.*

Por cada uno de los 4 bloques de contenidos de matemáticas, **hay una sección llamada MATES+MAGIA,** en la que aparecen divertidos trucos de magia que los alumnos pueden reproducir con sus compañeros. Estos trucos vienen acompañados de la explicación matemática correspondiente.

Para terminar, las soluciones de la autoevaluación se encuentran disponibles al final del libro.

Tabla 8

Actividades finales en el libro SAVIA SM

Actividades al final de la unidad		**Nº de actividades (nivel)**	**Contextualizado**	**Total**
Ejercicios para practicar.	**Conceptos elementales de estadística.**	4 (1)	4	4
	Gráficos estadísticos.	5 (1)	5	5
	Medidas de centralización.	5 (1)	5	5
	Medidas de dispersión.	2 (1)	2	4
		1 (2)	1	
		1 resuelta	1	
	Distribuciones bidimensionales.	4 (1)	2	10
		4 (2)	3	
		2 resueltas	0	
Problemas para resolver.		4 (2)	4	4
Actividades para pensar más. * encontrar el error.		4 (3)	3	5
		1 (2) *	1	
Problemas y actividades PISA.		1 resuelta	1	3
		2	2	
Autoevaluación.		4	3	4
				TOTAL = 44

Recapitulando lo visto hasta ahora, si hacemos un análisis cuantitativo, se observa que el número total de actividades es menor en Edelvives que en las otras dos editoriales.

Desde un punto de vista cualitativo, si consideramos que Edelvives también presenta un mayor número de actividades repetitivas y descontextualizadas, se puede afirmar que los recursos didácticos y aportaciones de esta editorial son inferiores en comparación con las otras dos:

- Edelvives tiene 39 actividades contextualizadas de un total de 62;
- Oxford Educación tiene 55 actividades contextualizadas de un total de 72;
- Savia SM tiene 63 actividades contextualizadas de un total de 71.

Savia SM es tan amplia en contenidos y conceptos estadísticos como Oxford Educación, lo que reafirma la limitación didáctica de Edelvives. Por ejemplo, los diagramas de cajas y bigotes junto con los rangos intercuartílicos no aparecen en Edelvives. Pero todas las diferencias y discrepancias existentes entre estos libros las pongo de manifiesto en un análisis comparativo a continuación.

2. Análisis comparativo de los tres libros de texto

Para profundizar más en el tratamiento que dan los libros de texto a la estadística en este análisis enumero las diferencias y discrepancias que he encontrado al indagar más en el contenido matemático que presentan.

Los aspectos que he tenido en cuenta son las diferencias que he encontrado en:

- los contenidos de los conceptos básicos y gráficos estadísticos (si aparecen o no en el libro, y con su correspondiente definición);
- la simbología matemática que utiliza para definir y manejar cada objeto matemático, es decir, el lenguaje estadístico que involucra letras (x, y, f, F, H, CV, s, r…), el símbolo sumatorio Σ, subíndices (x, y, i, j) y el símbolo promedio ($\bar{x}, \bar{y}$) entre otros;
- las fórmulas, comprobando si aparecen desarrolladas y justificadas debidamente.

Todas estas diferencias se sintetizan de forma ordenada y clasificada en las tablas 9, 10, 11, 12 y 13.

En ellas se recoge toda la información de una forma que es muy fácil comparar cada objeto matemático en los tres libros de texto. Dada la gran información que almacenan me limitaré a hacer una valoración que concentra las principales diferencias y conclusiones que se extraen de ellas.

De la tabla 9 se observa inmediatamente en Edelvives la ausencia de simbología matemática y de la definición de algunos conceptos básicos estadísticos tales como población o muestra. He comentado anteriormente la limitación de este libro pero sin detallar como ahora. Sin embargo, Oxford Educación y Savia SM tienen un lenguaje estadístico análogo aunque Oxford Educación evita el uso del sumatorio y las sumas aparecen desarrolladas.

Si nos fijamos en la tabla 9, el concepto de muestreo probabilístico aparece en Edelvives, no aparece en Oxford Educación y en Savia SM solo aparece el muestreo probabilístico aleatorio y estratificado. Además, el muestreo no probabilístico solo se encuentra en Edelvives. **Esto llama la atención, cabría preguntarse si las editoriales,**

conociendo el currículo de cursos anteriores, no han considerado necesario incluir ciertos conceptos en este curso ya que se han visto anteriormente.

Tabla 9

Diferencias de contenidos en los conceptos básicos de estadística

Conceptos básicos	Libro 1 **Edelvives**	Libro 2 **Oxford Educación**	Libro 3 **Savia SM**
Población.	Sin definición.	Sí.	Sí.
Muestra estadística.	Sin definición.	Sí.	Sí.
Muestreo probabilístico. (aleatorio)	Aleatorio simple. Sistemático. Estratificado.	No.	Aleatorio. Estratificado.
Muestreo **no probabilístico.** (no aleatorio)	De selección. Casuales. De bola de nieve.	No.	No.
Tipos de variables estadísticas.	No explica.	Sí, todas.	Sí, todas.
Variable estadística.	x_i	x_i	x_i
Marca de clase.	c_i	Marca de clase.	Marca de clase, x_i .
Nº total de datos.	Número total de datos, N.	$N = f_1 + f_2 + f_3 \dots$	$N = \sum_{i=1}^{n} f_i$
Frecuencia absoluta.	n_i	f_i	f_i
Frecuencia relativa.	$f_i = \frac{n_i}{N}$	$h_i = \frac{f_i}{N}$	$h_i = \frac{f_i}{N}$
Frecuencia relativa (%).	Porcentaje, p_i .	No	No
Frecuencia absoluta acumulada.	N_i	$F_i = f_1 + f_2 \dots + f_i$	$F_i = \sum_{j \le i} f_j$
Frecuencia relativa acumulada.	No.	$H_i = h_1 + h_2 \dots + h_i$	$H_i = \frac{F_i}{N} = \sum_{j \le i} h_j$
Frecuencia relativa acumulada (%).	No.	Porcentaje % .	$H_i(\%)$

Analizando los tipos de gráficos estadísticos que aparecen, como diferencias acusadas comentar que Edelvives no incluye en el tratamiento de datos la representación gráfica del polígono de frecuencias, ni en los diagramas de barras ni en los histogramas, pero sí explica cómo darse cuenta de las manipulaciones que presentan algunos gráficos para evitar caer en falacias interpretativas. **Aunque Savia SM no explica dichas falacias, es el libro que más tipos de gráficos estadísticos describe (ver tabla 10).**

Tabla 10

Diferencias de los tipos de gráficos estadísticos

Gráficos estadísticos	Libro 1 **Edelvives**	Libro 2 **Oxford Educación**	Libro 3 **Savia SM**
Diagrama de barras y polígono de frecuencias.	Diagrama sí. Polígono no.	Sí.	Sí.
Polígono de frecuencias acumuladas.	No.	Sí.	Sí.
Diagrama de sectores.	Sí.	Sí.	Sí.
Amplitud de cada sector (°).	Solo definición.	$h_n.360°$	$\frac{f_i}{N}.360° = h_i.360°$
Diagramas lineales.	No.	No.	Sí.
Histograma. Polígono de frecuencias.	Sí. Polígono no.	Sí.	Sí.
Explica falacias.	Sí.	No.	No.

De la tabla 11 destacar que en Edelvives los diagramas de cajas y bigotes no aparecen y los cuartiles solo los nombra sin dar una definición. Todo lo referente a las medidas de centralización y de posición aparece en los tres libros de forma análoga, si bien señalar que Edelvives denota con una "n" a la frecuencia absoluta de las variables estadísticas.

Tabla 11

Diferencias en los contenidos de los parámetros estadísticos unidimensionales: medidas de centralización y de posición

Medidas de centralización y de posición	Libro 1 **Edelvives**	Libro 2 **Oxford Educación**	Libro 3 **Savia SM**
Moda, Mo.	Sí.	Sí.	Sí.
Media, $\bar{x}$.	$\bar{x} = \frac{x_1 . n_1 + \cdots + x_n . n_n}{n_1 + \cdots + n_n} = \frac{\sum_{i=1}^{n} x_i . n_i}{N}$ Desarrollada.	$\bar{x} = \frac{x_1 . f_1 + \cdots + x_n . f_n}{N}$ $\bar{x} = \frac{\sum_{i=1}^{n} x_i . f_i}{N}$ Desarrollada.	$\bar{x} = \frac{x_1 + \cdots + x_n}{N} = \frac{\sum x_i}{N}$ $\bar{x} = \frac{x_1 f_1 + \cdots + x_n f_n}{N} = \frac{\sum x_i f_i}{N} = \Sigma x_i h_i$ Desarrollada.
Mediana, Me.	Sí.	Sí.	Sí.
Cuartiles Q_1, Q_2 y Q_3 , $Q_2 = $ Me .	Sin definir.	Sí.	Sí.

Las principales diferencias en las medidas de dispersión son cinco y residen en (ver tabla 12):

- **el recorrido intercuartílico** que solo aparece en el Savia SM;
- **la desviación media**, concepto que solo aparece en Oxford Educación;
- **la denotación de la varianza (V, σ^2, s^2) y de la desviación típica** (S, σ, s), que es distinta en cada libro;
- **el desarrollo de las fórmulas** aunque los tres libros emplean el sumatorio para simbolizar los parámetros de dispersión;
- **la distribución normal**, concepto que solo aparece en Savia SM.

Tabla 12

Diferencias en los contenidos de los parámetros estadísticos unidimensionales: medidas de dispersión

Medidas de dispersión	Libro 1 **Edelvives**	Libro 2 **Oxford Educación**	Libro 3 **Savia SM**
Recorrido o rango, R.	Sí.	Sí.	Sí.
Recorrido intercuartílico.	No.	No.	Sí.
Desviación media.	No.	$DM = \frac{\sum_{i=1}^{n} \lvert x_i - \bar{x} \rvert f_i}{N}$ Desarrollada.	No.
Varianza.	$V = \frac{\sum_{i=1}^{n} d_i^2 . n_i}{N} = \frac{\sum_{i=1}^{n} x_i^2 . n_i}{N} - \bar{x}^2$ Sin desarrollar.	$\sigma^2 = \frac{\sum_{i=1}^{n} x_i^2 f_i}{N} - \bar{x}^2$ Desarrollada.	$s^2 = \frac{\Sigma(x_i - \bar{x})^2 f_i}{N} = \frac{\Sigma x_i^2 f_i}{N} - \bar{x}^2$ Desarrollada.
Desviación típica.	$S = +\sqrt{V}$ Sin desarrollar.	$\sigma = \sqrt{\sigma^2}$ Sin desarrollar.	$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2 f_i}{N}} = \sqrt{\frac{\Sigma x_i^2 f_i}{N} - \bar{x}^2}$ Desarrollada.
Coeficiente de variación.	$CV = \frac{S}{\bar{x}}$	$CV = \frac{\sigma}{\bar{x}}$	$CV = \frac{s}{\bar{x}}$
Distribución normal.	No.	No.	Sí.

Y por último los parámetros de las distribuciones bidimensionales, contenido estadístico novedoso respecto a los cursos anteriores de la ESO, que se caracteriza por el estudio correlativo de las variables. **Como curiosidad significativa que la recta de regresión no se considere como un concepto estadístico relevante en Oxford**

Educación y solo se le dedique un ejercicio al final del libro en una sección ampliación llamada "Avanza" **(ver tabla 13).**

Tabla 13

Diferencias en los contenidos de los parámetros estadísticos bidimensionales

Distribuciones bidimensionales	Libro 1 **Edelvives**	Libro 2 **Oxford Educación**	Libro 3 **Savia SM**
Centro de gravedad.	No introduce el concepto como tal.	$(\bar{x}, \bar{y})$ Sin desarrollar.	$(\bar{x} = \frac{\sum x_i}{N}, \bar{y} = \frac{\sum y_i}{N})$ Desarrolladas.
Covarianza.	$S_{xy} = \frac{\sum_{i,j=1}^{k} n_{ij} \cdot x_i \cdot y_j}{N} - \bar{x}.\bar{y}$ Sin desarrollar.	$\sigma_{xy} = \frac{\sum_{i=1}^{n} x_i y_i f_i}{N} - \bar{x}.\bar{y}$ Sin desarrollar.	$S_{xy} = \frac{\Sigma(x_i - \bar{x}).(y_j - \bar{y})}{N} = \frac{\Sigma x_i y_j}{N} - \overline{xy}$ Sin desarrollar.
Coeficiente de correlación lineal de Pearson.	$r = \frac{S_{xy}}{S_x.S_y}$ Toma valores en el intervalo [-1,1].	$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$, Con $-1 \leq r \leq 1$.	$r = \frac{s_{xy}}{s_x s_y}$ Adimensional. Comprendido entre 1 y -1.
Recta de regresión lineal	$y - \bar{y} = \frac{S_{xy}}{S_x^2}(x - \bar{x})$; $m_x = \frac{S_{xy}}{S_x^2}$; $x - \bar{x} = \frac{S_{xy}}{S_y^2}(y - \bar{y})$	Aparece al final de la unidad en la sección "Avanza" pero no en las lecciones de contenido teórico	$y - \bar{y} = \frac{S_{xy}}{S_x^2}(x - \bar{x})$

A la vista de los resultados obtenidos del análisis comparativo de los tres libros, he llegado a la conclusión de que el libro Savia SM es el más completo debido a que:

- El rigor matemático y la extensión en la justificación desarrollada de las fórmulas del Savia SM sobresalen por encima de las otras editoriales, lo que le da un aspecto más sofisticado y avanzado matemáticamente hablando.
- Presenta más contenidos estadísticos y selecciona los más relevantes (por ejemplo, obvia la desviación media que apenas aporta información en el tratamiento estadístico pero sin embargo explica la interpretación conjunta de la media y la desviación típica en el modelo de la distribución normal).
- Todos los contenidos estadísticos están definidos y ejemplificados.
- Ofrece más recursos didácticos (digitales de ampliación, uso de la calculadora y GeoGebra).
- Dispone de un número mayor de actividades contextualizadas en las que:
    - Se modelizan fenómenos reales con variables estadísticas.
    - Se pide hacer inferencias al resolver los problemas.
- En las actividades resueltas y en los ejemplos desarrollados se explica el sentido estadístico de cada parámetro, dando una interpretación de los resultados.

C. Sobre los conocimientos previos: la estadística a lo largo de la educación obligatoria de un alumno

En España hace años que la Estadística está incluida en el currículo de la educación obligatoria, tanto en secundaria como en primaria, dentro de la asignatura de matemáticas. El alumno necesita una serie de conocimientos previos para afrontar el aprendizaje estadístico en 4º de ESO, por ello se debe indagar en cuáles son los contenidos que prescribe el currículum en cursos anteriores, es decir, buscar los referentes teóricos y antecedentes que posee en estadística al llegar a este curso.

A nivel estatal el Real Decreto 126/2014, de 28 de febrero, establece el currículo básico de la Educación Primaria.

A nivel autonómico el currículo de la Educación Primaria así como su aplicación en los centros docentes de la Comunidad Autónoma de Aragón se aprueba y autoriza en la ORDEN de 16 de junio de 2014, modificada por la ORDEN ECD/850/2016, de 29 de julio de la Consejera de Educación, Universidad, Cultura y Deporte.

Los contenidos del bloque de estadística en el currículo oficial de primaria *son los siguientes:*

- ***1º primaria:*** Recogida y recuento de datos en situaciones de observación. Registro e interpretación de datos en **pictogramas**.
- ***2º primaria:*** Recogida en situaciones de observación, recuento y agrupación de datos en función de un criterio. Registro e interpretación de gráficos sencillos **(diagramas de barras y pictogramas).**
- ***3º primaria:*** Recogida, ordenación y clasificación de datos en función de un criterio. Realización e interpretación de gráficos sencillos **(diagramas de barras y circulares).**
- ***4º primaria:*** Recogida, ordenación y clasificación de datos en función de más de un criterio. Realización e interpretación de gráficos sencillos **(diagramas de barras, lineales y circulares).**
- ***5º primaria:*** Recogida y clasificación de datos cualitativos y cuantitativos. Construcción de **tablas de frecuencias absolutas.** Iniciación intuitiva a los conceptos de **media aritmética, rango, frecuencia y moda.** Realización e

interpretación de gráficos sencillos **(diagramas de barras, lineales, circulares…).**

- ***6º primaria:*** Recogida y clasificación de datos cualitativos y cuantitativos. Construcción de **tablas de frecuencias absolutas y relativas.** Iniciación intuitiva a los conceptos de **media aritmética, rango, frecuencia y moda.** Realización e interpretación de gráficos sencillos: **diagramas de barras, poligonales y sectoriales.** Análisis crítico de las informaciones que se presentan mediante gráficos estadísticos.

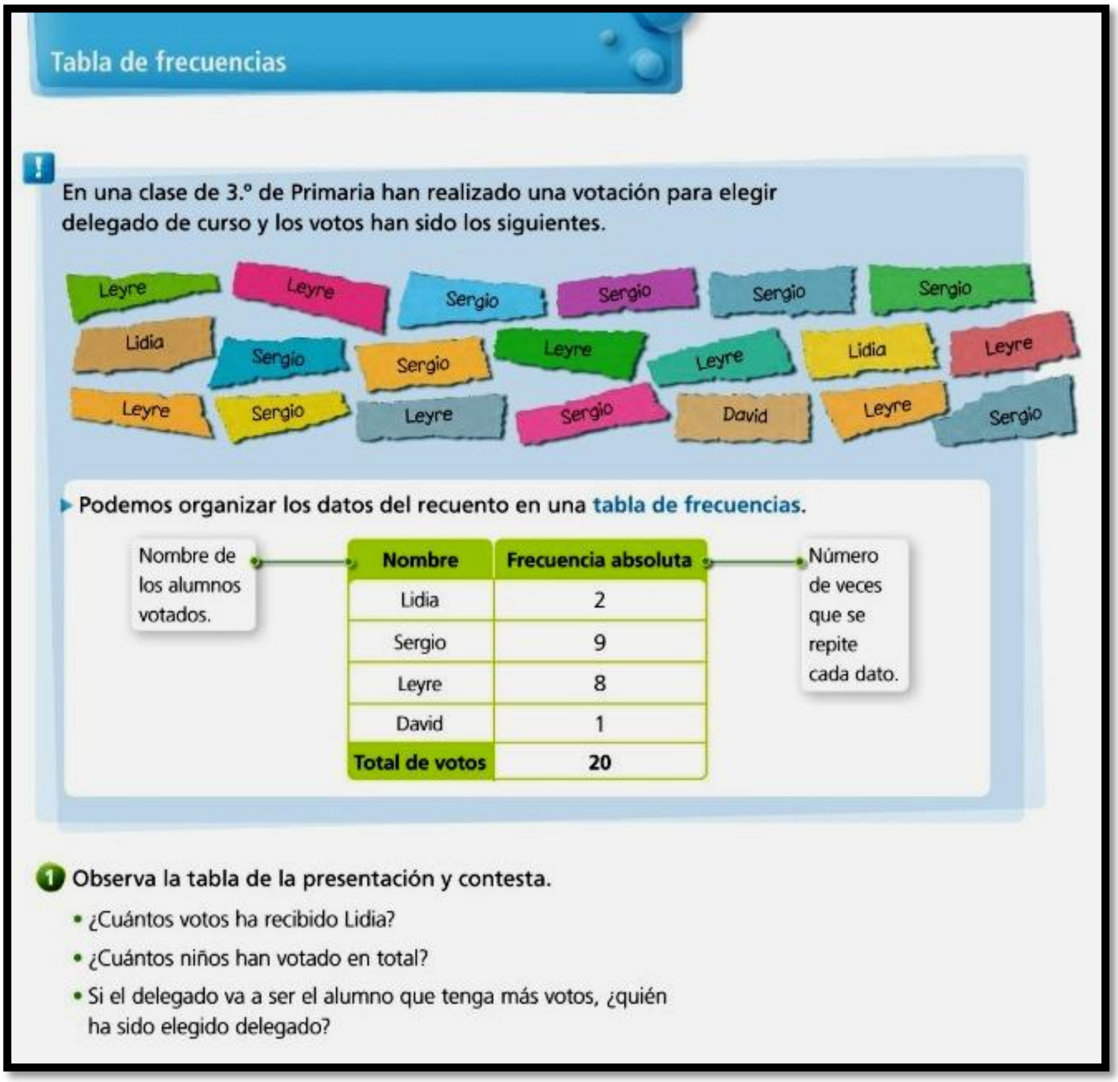


Tabla de frecuencias

En una clase de 3.º de Primaria han realizado una votación para elegir delegado de curso y los votos han sido los siguientes.

Leyre, Leyre, Sergio, Sergio, Sergio, Sergio, Lidia, Sergio, Sergio, Leyre, Leyre, Lidia, Leyre, Leyre, Sergio, Leyre, Sergio, David, Leyre, Sergio

Podemos organizar los datos del recuento en una tabla de frecuencias.

Nombre de los alumnos votados.

Nombre	Frecuencia absoluta
Lidia	2
Sergio	9
Leyre	8
David	1
Total de votos	**20**

Número de veces que se repite cada dato.

1 Observa la tabla de la presentación y contesta.

- ¿Cuántos votos ha recibido Lidia?
- ¿Cuántos niños han votado en total?
- Si el delegado va a ser el alumno que tenga más votos, ¿quién ha sido elegido delegado?

Figura 10. Problema de estadística en 3º de Primaria del libro EDELVIVES.

Como ya he citado anteriormente, a nivel estatal el currículo básico de secundaria está marcado por la Ley Orgánica 2/2006, de 3 de mayo, de Educación (LOE) modificada por la Ley Orgánica 8/2013, de 9 de diciembre, para la Mejora de la Calidad Educativa (LOMCE).

Y a nivel autonómico el currículo oficial de matemáticas se apoya en la orden ECD/489/2016, de 26 de mayo.

***Los contenidos del bloque de estadística en el currículo oficial de secundaria** son los siguientes:*

- ***1º ESO:*** Población e individuo. Muestra. Variables estadísticas. Variables cualitativas y cuantitativas. Frecuencias absolutas y relativas. Organización en tablas de datos recogidos en una experiencia. Diagramas de barras, y de sectores. Polígonos de frecuencias. Medidas de tendencia central.
- ***2º ESO:*** Población e individuo. Muestra. Variables estadísticas. Variables cualitativas y cuantitativas. Frecuencias absolutas y relativas. Organización en tablas de datos recogidos en una experiencia. Diagramas de barras, y de sectores. Polígonos de frecuencias. Medidas de tendencia central. Medidas de dispersión.
- ***3º ESO de matemáticas académicas***: Fases y tareas de un estudio estadístico. Población, muestra. Variables estadísticas: cualitativas, discretas y continuas. Métodos de selección de una muestra estadística. Representatividad de una muestra. Frecuencias absolutas, relativas y acumuladas. Agrupación de datos en intervalos. Gráficas estadísticas. Parámetros de posición. Cálculo, interpretación y propiedades. Parámetros de dispersión. Diagrama de caja y bigotes. Interpretación conjunta de la media y la desviación típica.

Después de hacer está revisión de los contenidos estadísticos de cursos anteriores, se puede afirmar que el alumno posee los conocimientos y recursos necesarios para iniciar la estadística de 4º de ESO.

Por otro lado hay que tener en cuenta la evolución de la materia en el propio año académico, ya que la estadística supone el último bloque de materia. Las nociones adquiridas secuencialmente durante el curso son igualmente importantes y suponen la base sobre la que se construye el aprendizaje estadístico:

- los porcentajes;
- el lenguaje algebraico;
- el lenguaje gráfico para funciones;
- los intervalos;
- el concepto de frecuencia en situaciones donde interviene el azar.

1. Evaluación inicial de conocimientos previos

Para atraer la atención del alumnado, el inicio de la estadística en el aula se realizará mediante la reproducción de un vídeo corto sobre la historia de la estadística. Este vídeo se encuentra en los recursos didácticos de la Red Educativa Digital Descartes y es el siguiente:

Figura 11. Vídeo historia de la estadística https://youtu.be/lTnJLGhmW8k .

Una vez acabado este vídeo de introducción, se llevará a la práctica un breve cuestionario para repasar y evaluar los conocimientos previos de los alumnos. Esta evaluación inicial involucra una serie de actividades sencillas como las que propongo a continuación.

1. Clasifica cada una de las siguientes variables en cualitativas, discreta o continua, escribiendo una X en el recuadro correspondiente (tabla 14):

Tabla 14

Tabla para clasificar tipos de variables estadísticas

VARIABLES	CUALITATIVA	DISCRETA	CONTINUA
Tipo de música preferida.			
Páginas de un libro.			
Estatura.			
Asignaturas pendientes.			
Nº de hijos.			
Longitud del pie.			

2. Calcula los siguientes porcentajes.

a) El 75% de 80.
b) El 65% de 130.
c) El 2% de 75.
d) El 25% de 32.
e) El 10% de 37.

3. En un grupo de 20 niños se ha tomado el peso en kg de cada uno de ellos y se ha obtenido el siguiente resultado:

41, 31, 35, 31, 32   34, 33, 39, 36, 39
38, 35, 31, 34, 40   37, 32, 35, 33, 31

a) Ordena los datos en una tabla de frecuencias.
b) Ahora agrupa los datos en intervalos con la misma amplitud (de 3 en 3 kg) y construye la tabla de frecuencias correspondiente.

4. Las edades de un grupo de 9 amigas son: 12, 14, 13, 16, 13, 15, 15, 17 y 13. Calcula las medidas de centralización (media, moda y mediana).

D. Sobre la razón de ser de la introducción escolar de la estadística

Según Holmes (2002), la enseñanza de la estadística y probabilidad se inició en 1961 en el currículo de Inglaterra en forma opcional para los estudiantes de 16 a 19 años que querían especializarse en matemáticas. El motivo era mostrar su aplicación en otras materias.

Actualmente el objetivo principal de introducir el conocimiento estadístico en la formación básica de los ciudadanos es proporcionar una cultura estadística para que se puedan desenvolver en la sociedad. En la literatura destaca el modelo de Gal (2002) que describe la alfabetización estadística mediante dos componentes básicos interrelacionados:

> *a)"capacidad para interpretar y evaluar críticamente la información estadística, los argumentos apoyados en datos o los fenómenos estocásticos que las personas pueden encontrar en diversos contextos, incluyendo los medios de comunicación, pero no limitándose a ellos, y b) capacidad para discutir o comunicar sus opiniones respecto a tales informaciones estadísticas cuando sea relevante"* (Gal, 2002, pp. 2-3).

Sin embargo, aunque la estadística se incluye en el currículo oficial esto no implica que se enseñe, muchas veces se omite porque se deja para el final del temario o se dispone de poco tiempo para introducirla adecuadamente. La razón reside en que su experiencia y los contenidos curriculares no han abordado el tratamiento de la estadística con un adecuado tratamiento didáctico.

En el proceso de asimilación de los conceptos estadísticos más complejos se debe valorar la dificultad que tienen los alumnos en el aprendizaje, el enfoque que dan los libros de texto a la estadística, el rápido avance de los medios tecnológicos y científicos, y la formación del profesorado acorde a estos cambios.

Tradicionalmente la estadística ha priorizado el cálculo y la demostración de las propiedades estadísticas, aspectos que ahora pierden importancia debido a la aplicación de las nuevas tecnologías, como por ejemplo la hoja de cálculo. Por ello ahora la atención debe centrarse en:

- **la interpretación y significado** de las tablas y los gráficos obtenidos por medio de las herramientas tecnológicas;
- **en fomentar el espíritu crítico y creativo** de los alumnos ya que deben extraer y razonar sus propias conclusiones;
- **en su marcado carácter interdisciplinar** ya que se aplica en estudios sociales y científico-tecnológicos, estudios que deben ser cercanos a los intereses del alumnado.

1. Origen y evolución histórica de la estadística

Estadística proviene del término alemán Statistik, acuñado por el profesor Gottfried Achenwall (1719 – 1772), que significa ciencia del Estado porque su origen reside en la necesidad de hacer estudios descriptivos sobre todos los integrantes de una población.

Figura 12. Retrato de Gottfried Achenwall.

La razón de ser de la introducción de la estadística en el ámbito escolar nada tiene que con las razones históricas que dieron lugar a sus orígenes y cuyas evidencias de existencia se sitúan cronológicamente en la **Prehistoria y en las civilizaciones antiguas, es decir, en los comienzos de la humanidad.**

- **En el comienzo de la historia** se utilizaban muescas en piedra y huesos, señales en pieles, palos y paredes de cuevas para representar y contar personas, animales o pertenencias.
- **Los NURAGAS**, en la isla de Cerdeña (años 1800-800 a. de C.), donde aparecen indicios sobre el control de sus posesiones.
- **Los BABILÓNICOS** hacia el año 3000 a. de C. ya usaban pequeñas tablillas de arcilla para anotar en escritura cuneiforme datos relacionados con la producción agrícola y ganadera.
- **Los EGIPCIOS** en el año 3050 a. de C. analizaban los datos de la población y la renta del país con el propósito de construir las pirámides.
- **En CHINA** hacia el año 2.200 a. de C. bajo el mandato del rey Yao (Tao) hicieron una estadística agrícola, industrial y comercial
- **En la ANTIGUA GRECIA** hicieron importantes observaciones estadísticas en censos con fines tributarios, sociales y militares.
- **Los ROMANOS** más tarde, 555 a. de C., maestros de la organización política fueron los que mejor supieron emplear los recursos en la estadística: nacimientos, defunciones y matrimonios, recuentos periódicos de ganado, riquezas contenidas en las tierras conquistadas, etc.

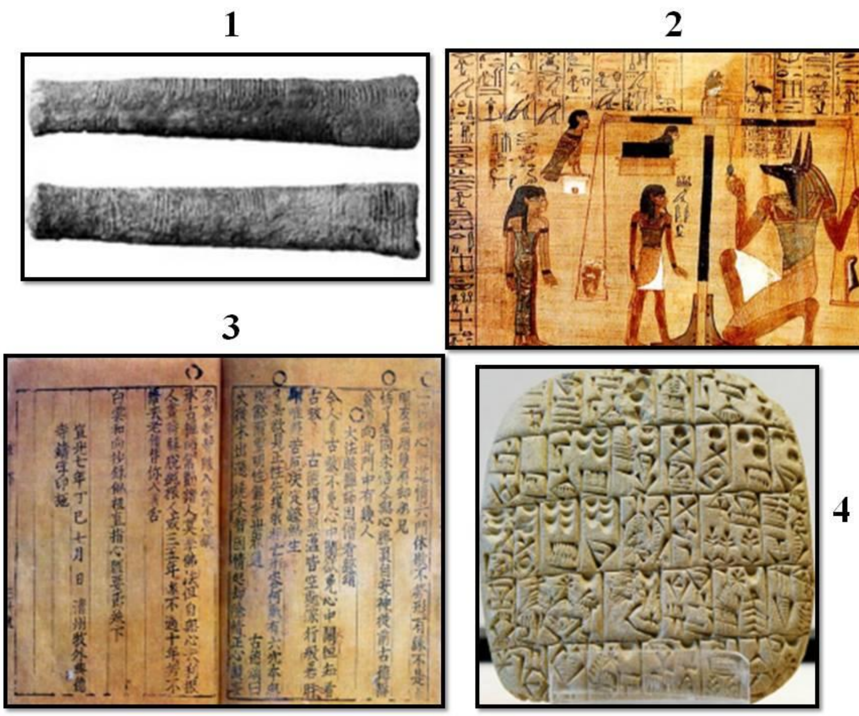


Figura 13. Evidencias de la estadística en la Prehistoria y las civilizaciones antiguas: 1- muescas hechas en piedra, 2- papiro egipcio, 3-papiro chino, 4- tablilla babilónica con escritura cuneiforme.

Este tipo de estudios continuaron evolucionando y desarrollándose a lo largo de la Historia hasta la actualidad. La actividad estadística durante la **Edad Antigua** consistía principalmente en:

- elaborar censos de población, catastros para facilitar la gestión tributaria y registros para organizar el servicio militar;
- establecer repartos de tierras u otros bienes materiales;
- determinar el derecho al voto de los ciudadanos.

A lo largo de la **Edad Media** su finalidad no experimentó grandes avances y se mantuvo reducida a recopilar información con carácter exclusivamente organizativo. El brote de peste que se produce en Inglaterra en la década de 1500 provoca un cambio al publicarse semanalmente datos sobre defunciones que posteriormente se añadieron a los datos de nacimientos por sexo, realizándose predicciones sobre los mismos.

Es entonces durante la **Edad Moderna** que se comienza a hacer análisis de datos para sacar conclusiones y su uso se amplía. El desarrollo científico-matemático que se dio en este periodo aportó mucho a la estadística: por un lado, **importantes científicos de la época como Copérnico, Galileo, Bacon, Descartes, etc. desarrollan el método que debe seguir un estudio para ser considerado científico, y por otro lado, los matemáticos Pascal y Fermat establecen las bases de la Teoría de la Probabilidad que estudia los fenómenos aleatorios.**

La estadística continúa evolucionando cada vez más deprisa durante la **Edad Contemporánea**, época donde los estudios demográficos, económicos y sociales adquieren cada vez más importancia. El desarrollo de las Matemáticas y de otras ciencias proporciona técnicas analíticas que hacen posible establecer relaciones entre las variables y hacer predicciones. El desarrollo del muestreo y de la inferencia estadística permite estudiar la población a través de sólo una parte de ella lo que facilita la recogida y el procesamiento de datos. **De los trabajos científicos como Laplace, Gauss y Legendre surge la teoría de los errores de observación y el método de mínimos cuadrados. Y de las investigaciones de Galton y Pearson surgen los conceptos de correlación y curva de regresión.** Estas nuevas técnicas son muy aplicadas en estadística actualmente.

La máquina de Galton es un dispositivo diseñado por **Francis Galton** que enseña como las colisiones al azar de unas bolitas que caen alrededor de unos clavos de madera sobre unos casilleros forman lo que en estadística se conoce como una distribución normal o campana de Gauss. Los clavos de madera que hay en el camino están uniformemente repartidos, lo que hace que al rebotar contra ellos las bolitas puedan ir a izquierda o derecha, completamente al azar siempre que el tablero se mantenga equilibrado. Lo que sucede es que hay más maneras de llegar al centro que a los lados y en vez de una distribución uniforme, surge una peculiar forma acampanada.

Figura 14. Máquina de Galton.

A finales del siglo XX se extiende el uso de herramientas tecnológicas para tratamiento estadístico de datos entre los que destacan:

- **la hoja de cálculo Excel/Calc** muy útil cuando no se tratan muchos datos;
- **el programa estadístico SSPS** muy usado en las ciencias sociales y aplicadas, además de las empresas de investigación de mercado cuyo nombre corresponde a las siglas del acrónimo *Statistical Package for the Social Sciences*;
- otros programas importantes a nivel mundial como **S-Plus, MiniTab, Statgraphics, StatSoft y PH-Stat.**

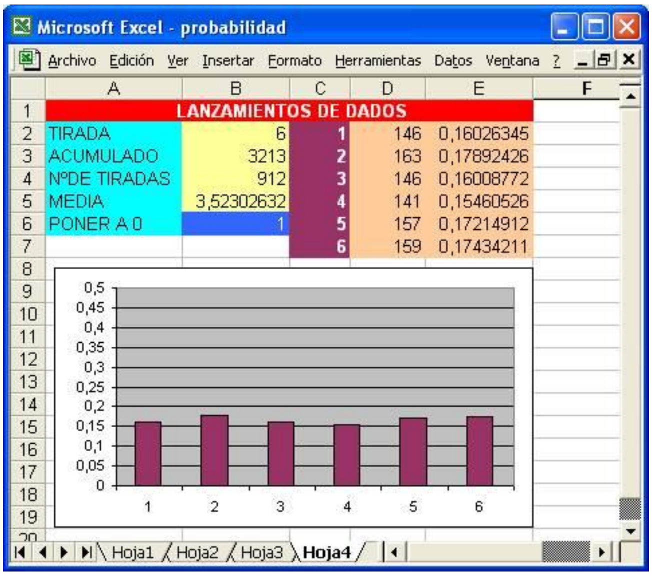


Figura 15. Hoja de cálculo EXCEL.

En el siglo XXI se produce el auge del "Big Data" que estudia datos complejos que llegan en volúmenes enormes a una velocidad cada vez mayor, con múltiples fuentes autónomas y en distintos soportes. Estos datos no se pueden procesar mediante un ordenador de capacidad tradicional y la metodología debe adaptarse a esta situación de información masiva mediante las nuevas técnicas estadísticas de Big Data.

Figura 16. Las nuevas técnicas estadísticas de BIG DATA para datos masivos.

Todos los países disponen de oficinas de estadística que trabajan en registrar los datos estadísticos oficiales: sobre población, educación, productos de consumo, sanidad, tasa de paro, actividad económica, etc. En España hay varios organismos que hacen estadísticas nacionales y uno de ellos es el INE (Instituto Nacional de Estadística).

2. Problemas de razón de ser y metodología para su implementación

A diario nos encontramos una gran cantidad de datos cuando vemos la televisión, navegamos por internet o leemos cualquier tipo de periódico, especialmente los deportivos. **Para poder comprender la información que nos proporcionan estos datos en gráficos o tablas, inferir conclusiones o hacer predicciones basadas en la interpretación que deducimos de ellos, se debe poner en funcionamiento nuestras capacidades y habilidades adquiridas en el aprendizaje estadístico. Por esta razón, lo lógico sería comenzar con problemas que den este sentido al estudio de la estadística y que constituyan una razón de ser para su implementación en el aula.** Un ejemplo de problema se iniciación al sentido estadístico podría ser el que enuncio y desarrollo a continuación en varios apartados.

<u>PROBLEMA</u>: "UN ESTUDIO SOBRE EL DEPORTE EN SECUNDARIA"

Vamos a analizar la práctica de deportes en un instituto de enseñanza secundaria por los alumnos y alumnas de toda la ESO.

<u>Apartado A</u>: **¿Qué deporte practican más los alumnos?**

En un instituto de secundaria se practican los siguientes deportes: fútbol, baloncesto, voleibol, ping-pong, atletismo y ajedrez. Se hace una encuesta censal entre todos los alumnos y alumnas de la ESO en la que se pregunta cuál es el deporte que practican con más frecuencia. Los alumnos solo pueden escoger un deporte al responder la encuesta.

Los resultados se recogen en la tabla 15.

Tabla 15

Resultados de la encuesta sobre el deporte que más practican los alumnos y alumnas

CURSO	1ºESO		2º ESO		3º ESO		4º ESO	
SEXO	M	F	M	F	M	F	M	F
Fútbol	12	0	13	0	8	0	8	0
Baloncesto	10	10	9	10	6	14	6	16
Voleibol	3	4	4	2	0	2	0	2
Ping-pong	2	1	2	2	0	2	2	2
Atletismo	9	5	8	6	8	7	8	8
Ajedrez	4	1	4	1	4	0	4	0
Ninguno	10	14	10	14	4	10	2	7

Para comprender la tabla estadística es preciso conocer el significado de cada número escrito en ella. Hay que distinguir tres conceptos: curso, sexo y deporte preferido. Por ejemplo, el 12 de la primera casilla representa el número de chicos de 1º de ESO que son aficionados al fútbol.

Pero, ¿qué información se puede obtener de esta tabla? Los alumnos tienen que deducir:

- el número total de alumnos que hay en un curso y en la ESO;
- cuántos alumnos practican preferentemente un deporte en cada curso o cuántos no practican deporte.

Para ello deben hacer sumas parciales y totales en función de los tres conceptos señalados anteriormente, por lo que para seguir con nuestro análisis y manejar más fácilmente los datos ampliaremos la tabla sumando todos los datos ordenadamente (ver tabla 16).

Tabla 16

Tabla ampliada de los resultados de la encuesta

CURSO	1ºESO			2º ESO			3º ESO			4º ESO			Totales		
SEXO	M	F	T	M	F	T	M	F	T	M	F	T	M	F	T
Fútbol	12	0		13	0		8	0		8	0				
Baloncesto	10	10		9	10		6	14		6	16				
Voleibol	3	4		4	2		0	2		0	2				
Ping-pong	2	1		2	2		0	2		2	2				
Atletismo	9	5		8	6		8	7		8	8				
Ajedrez	4	1		4	1		4	0		4	0				
Ninguno	10	14	24	10	14		4	10		2	7				
Totales	50	35	85										160	140	300

Los alumnos deberán completar la tabla ampliada comprobando las sumas que aparecen en rojo y completando las que faltan para poder seguir con nuestro análisis y responder a otras preguntas como:

- ¿qué proporción de chicos y de chicas hay en cada curso?;
- ¿qué porcentaje de ajedrecistas hay en cada curso?;
- ¿cómo evoluciona la abstención en el deporte al subir de curso?;
- ¿cómo evoluciona la participación femenina en atletismo al subir de curso?, ¿y en voleibol?, ¿qué deporte o deportes se practican más en cada curso?;
- según los datos de la encuesta, ¿encuentras algún detalle significativo en la práctica de algún deporte? Dime qué aspectos sociales implica para ti.

Las conclusiones que se extraen del ejercicio de este apartado son las siguientes:

- nos resulta más fácil manejar e interpretar la información extraída de la encuesta porque los resultados se presentan ordenados en una tabla;
- para calcular las proporciones y porcentajes se ha tenido que ampliar la tabla para añadir las sumas parciales y totales;
- los resultados de la tabla ampliada se pueden obtener sumando de dos formas distintas (por ejemplo, 85 se obtiene sumando 50 más 35 o sumando los números que hay encima de él).

Apartado B: **¿Cuántas horas semanales se dedica a cada deporte?**

A cada alumno de una clase de 4º de ESO se pregunta cuántas horas dedica semanalmente a practicar su deporte preferido. Los resultados de la encuesta son los siguientes:

3,2,4,5,4	0,3,4,5,0	3,6,2,4,5
3,4,3,3,4	2,2,4,2,0	2,4,2,7,5

- ¿Cuántos alumnos son en clase?
- Cuántas respuestas distintas dan los alumnos, ¿cuáles son?
- Cuántos alumnos dan cada respuesta.
- Haz una tabla que recoja de forma ordenada y clasificada todos estos datos. ¿Cómo lo harías?
- Representa los datos ordenadamente y agrupados por cada tipo de respuesta mediante un gráfico de barras donde la longitud de la barra sea proporcional al número de respuestas de cada tipo. Toma como referencia la tabla creada anteriormente.
- En función de los resultados obtenidos, ¿qué valor puede representar toda la distribución? De qué modo harías el cálculo de este valor y cómo lo indicarías en la gráfica anterior para poder visualizar la separación o diferencia entre cada dato y este valor.

En este apartado se modeliza una situación donde se introduce el concepto de frecuencia de una variable y su representación gráfica. Se promueve al alumno a investigar sobre los resultados para llegar al valor medio y la desviación típica, es decir, las medidas de centralización.

Apartado C: **Altura media de los jugadores de baloncesto**

El comité organizador de la liga masculina de baloncesto de 4º de ESO ha hecho un estudio de las medias de las tallas de todos los equipos. La media ha sido útil para un primer conocimiento de la altura de los equipos pero resulta insuficiente porque hay equipos que tienen la misma media pero sus distribuciones son muy diferentes. Por ello el comité añade a su estudio un parámetro llamado desviación típica que mide cómo se dispersan los datos de esta talla media. Aquí aparecen los datos de 2 equipos de la liga:

Tabla 17

Resultados del estudio de las tallas de los equipos A y B

Equipo	**A**	**B**
Media	169,16 cm	169,5 cm
Desviación típica	6,1	3

De este estudio se han obtenido los siguientes gráficos pero no sabemos su correspondencia (ver figura 15).

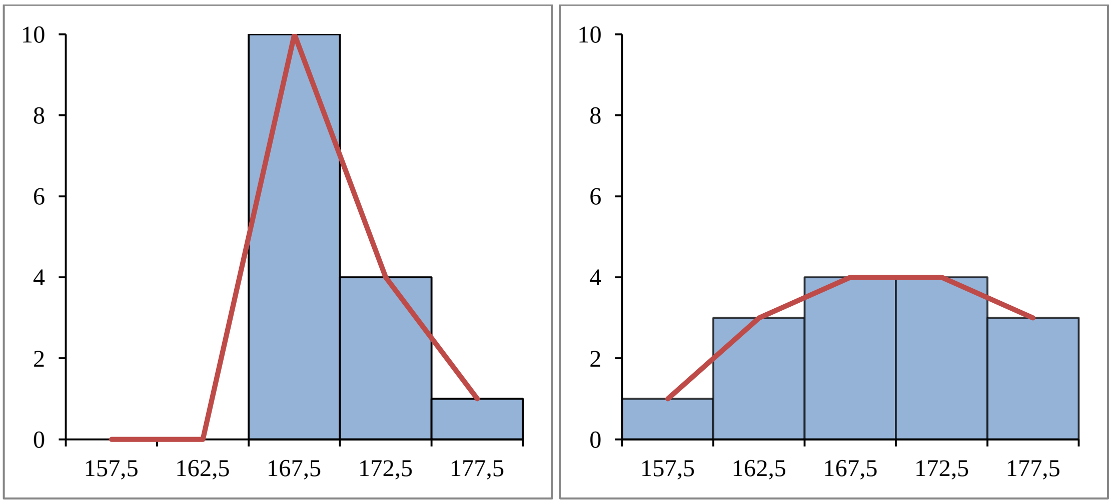


Figura 15. Gráficos del estudio de las tallas de los equipos A y B.

- ¿Cómo saber qué gráfica corresponde a cada par de parámetros?
- Explícame cómo interpretas ese polígono lineal que aparece en el gráfico.
- Si de un equipo sólo sabemos que su media es 171,5 y su desviación típica es 4,5: ¿podemos saber cuántos jugadores miden más de 175 cm?, ¿y cuántos miden menos de 165 cm? Explícame qué consecuencias cuantitativas se pueden deducir.

El objetivo de este apartado es desarrollar su espíritu crítico para interpretar correctamente las medidas de centralización y evitar o corregir posibles errores adquiridos.

<u>Apartado D</u>: **Rendimiento encestando de las jugadoras de baloncesto**

Cuatro jugadoras del equipo de baloncesto de 4º de ESO han hecho 10 lanzamientos a canasta a una distancia de 1m, otros 10 desde 2 m y así sucesivamente hasta 7 m. En cada caso se ha anotado el número de encestes:

Tabla 18

Resultados de los lanzamientos de las cuatro jugadoras de baloncesto

Distancia / **Jugadora**	1m	2m	3m	4m	5m	6m	7m	8m
Clara	9	10	6	4	2	0	1	
Eva	7	6	7	4	2	4	1	
Marta	3	4	0	1	0	2	1	
Sofía	10	8	9	9	6	7	4	

- A la vista de la tabla, ¿sabrías decir si alguna de las cuatro jugadoras es más eficaz en el enceste?
- Para estudiar la eficacia en el enceste debemos representar una distribución que asocie la distancia a la que se producen los lanzamientos y el número de canastas conseguidas. Representa en un gráfico estos pares de valores como puntos en un diagrama cartesiano para cada jugadora. ¿Cómo es la relación entre las dos variables observadas en cada gráfico? ¿Se observa algún tipo de tendencia?
- Compara los resultados para las cuatro jugadoras y responde a la primera pregunta ahora. ¿Podrías predecir el número de encestes a 8 m para cada jugadora?

En este último apartado consideramos la correlación entre dos variables observadas e induciremos al alumno a hacer un cálculo predictivo para una distancia de 8 m.

La metodología de esta actividad es combinar el trabajo individual de cada alumno en el desarrollo de cada apartado del problema con el grupal mediante un posterior debate para exponer brevemente las conclusiones que cada uno ha extraído de su ejercicio, tanto a nivel didáctico como aplicado a la vida diaria. Los aspectos estadísticos más importantes se van introduciendo progresivamente, por ello el orden de los apartados y los momentos de estudio son relevantes para poder hacer una adecuada construcción del conocimiento, así como del lenguaje matemático, validación e institucionalización del objeto matemático en cuestión.

E. Sobre el campo de problemas, técnicas y tecnologías: diseño de la secuencia didáctica

El diseño de la secuencia didáctica de basa en el aprendizaje a través de la resolución de problemas y comprende siete sesiones distribuidas del siguiente modo:

- **cuatro sesiones destinadas a la enseñanza-aprendizaje de la estadística unidimensional,** que supone un repaso y profundización de conceptos ya vistos en cursos anteriores;
- **y tres sesiones destinadas a la enseñanza-aprendizaje de la estadística bidimensional**, cuyos conceptos se introducen por primera vez en secundaria.

Durante la secuencia el proceso de evaluación será continuo y formativo. Al terminar la secuencia, se realizará una prueba escrita de lo que se ha visto en clase.

1. Primer campo de problemas: muestreo, tablas de datos y frecuencias (1 sesión)

<u>Problema 1.</u> (Savia SM – ejemplo 1)

Una multinacional textil quiere estudiar las tallas usadas por sus potenciales clientes. Sabe que el 80% son hombres y la cuarta parte son menores de 30 años. Las clientes femeninas menores de 30 años son el 60% del total. Van a tomar una muestra aleatoria estratificada de mil clientes, formada por:

	Total	< 30 años	> 30 años
Hombres	800	200	600
Mujeres	200	120	80

¿Es representativa esta muestra?

<u>Solución:</u>

Es representativa porque los individuos que forman la muestra están en la misma proporción que en la población.

<u>Problema 2.</u> (Oxford Educación – enunciado del problema 3)

En un centro escolar se va a escoger una muestra de 50 estudiantes para realizar una encuesta. Si el 58% son chicas, ¿cuántas chicas y cuántos chicos deberán formar la muestra para que esta sea representativa?

<u>Solución:</u>

Chicas	50 x 0,58 = 29	**58%**
Chicos	50 x 0,42 = 21	42%
Total	**50**	100%

<u>Problema 3.</u> (Oxford Educación – enunciado del problema 4)

Con el fin de estudiar la calidad de un producto envasado por dos marcas de conservas de bonito del norte, A y B, debemos seleccionar una muestra estratificada de 10 latas. Si disponemos de 54 latas de la marca A y de 36 de la marca B, ¿cuántas latas de cada marca deben incluirse en la muestra?

<u>Solución:</u>

	Nº de latas	Nº de latas (%)	Muestra
Marca A	**54**	60%	6
Marca B	**36**	40%	4
Total	90	100%	10

<u>Problema 4.</u> (Santillana – enunciado adaptado del problema 25)

Indica el tipo de variable estadística que estudiamos y razona, en cada caso, si sería mejor analizar una muestra o la población:

a) La altura y el peso de los alumnos de una clase.

b) La marca de los coches de una ciudad.

c) La edad de los habitantes de un país.

Solución:

a) Es una variable cuantitativa continua. La población son los alumnos de la clase, no es necesario tomar una muestra.

b) Es una variable cualitativa. La población son los coches de la ciudad y analizar todos los datos resulta costoso. En este caso es necesario seleccionar una parte de la población, es decir, una muestra representativa para realizar el estudio.

c) Es una variable cuantitativa discreta. La población son los habitantes de un país, se debería tomar un grupo de personas de la población elegidas al azar con objeto de poder extraer conclusiones válidas para el total de la misma.

Problema 5. (Oxford Educación – enunciado del problema 7)

En el centro médico de un colegio han realizado un estudio sobre el peso, en kilogramos, de 20 alumnos, y los resultados han sido:

61	**66**	**56**	**66**	**49**
56	**58**	**62**	**71**	**61**
63	**53**	**45**	**51**	**49**
57	**60**	**63**	**49**	**58**

Ordena estos datos en una tabla de frecuencias agrupándolos en cinco intervalos de la misma amplitud.

Solución:

Amplitud del intervalo → (Dato Máximo – Dato Mínimo): 5 = (71 – 45) / 5 = 5,2

Variable Peso en kg	Frecuencia absoluta
[45,51)	4
[51,57)	4
[57,63)	7
[63,69)	4
[69,75)	1

Problema 6. (Savia SM – enunciado del problema 3)

Los resultados de una encuesta a 40 jóvenes sobre el número de horas que utilizan una consola el fin de semana son:

4	**2**	**3**	**7**	**6**	**4**	**3**	**7**	**6**	**8**
3	**4**	**3**	**5**	**3**	**2**	**1**	**0**	**5**	**4**
3	**7**	**8**	**0**	**1**	**6**	**4**	**5**	**7**	**6**
1	**4**	**3**	**7**	**5**	**4**	**3**	**1**	**0**	**3**

a) ¿De qué tipo es la variable estadística?

b) Haz una tabla de frecuencias indicando la frecuencia absoluta, relativa, y las frecuencias acumuladas de cada dato.

c) ¿Qué porcentaje utiliza la consola menos de 3 horas? ¿Y más de 6?

Solución:

a) La variable estadística es cuantitativa discreta.

b)

Variable Nº de horas consola	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada	Frecuencia relativa acumulada (%)
0	3	0,075	3	0,075	7,5
1	4	0,1	7	0,175	17,5
2	2	0,05	9	0,225	22,5
3	9	0,225	18	0,45	45
4	7	0,175	25	0,625	62,5
5	4	0,1	29	0,725	72,5
6	4	0,1	33	0,825	82,5
7	5	0,125	38	0,95	95
8	2	0,05	**40**	**1**	**100**
	N = 40	**1**			

c) El porcentaje que utiliza la consola menos de 3 horas es 22,5 %.

El porcentaje que utiliza la consola más de 6 horas es 17,5% (=100% - 82,5%).

Problema 7. (Santillana – enunciado del problema 29)

Explica cómo completarías una tabla de frecuencias conociendo solo las frecuencias absolutas acumuladas. ¿Podrías hacer lo mismo con las frecuencias relativas acumuladas?

Solución:

Conociendo las frecuencias absolutas acumuladas, sabemos cuál es el nº total de datos N, que es la mayor frecuencia acumulada. Podemos calcular las frecuencias absolutas restando a la mayor frecuencia acumulada, la anterior frecuencia acumulada y así sucesivamente.

Una vez obtenidas las frecuencias absolutas, y conocido N, se pueden calcular las frecuencias relativas, las frecuencias relativas acumuladas y las frecuencias relativas acumuladas en %. De este modo se completa la tabla de frecuencias.

Si no se conoce el número total de datos N, sabiendo solamente las frecuencias relativas acumuladas no se puede completar la tabla de frecuencias.

Técnicas o modificaciones de una técnica que se ejercitan

Las técnicas para resolver estos problemas involucran las tablas de datos y las tablas de frecuencias de modo que se ejercita:

- la ordenación numérica de los datos (de menor a mayor) teniendo en cuenta la frecuencia con que aparecen;
- el cálculo de porcentajes y en sentido de las proporciones;
- la proporción en relación a la unidad;
- el sentido y razonamiento estadístico de muestra y proporción;
- la división de una secuencia numérica en intervalos semiabiertos por la derecha.

Tecnologías y proceso de institucionalización de los distintos aspectos del objeto matemático

Las tecnologías que se repasan o adquieren en esta sesión son:

- la diferencia entre muestra y población;
- qué es una muestra representativa;
- saber diferenciar los tipos de variables estadísticas;
- la diferencia entre los conceptos “discreto” y “continuo”;
- el concepto de frecuencia de los datos;
- el análisis y tratamiento de los datos mediante tablas.

Este campo de problemas supone un reencuentro con técnicas aprendidas anteriormente, de modo que posee ya las herramientas necesarias, tras una primera exploración de los problemas, para constituir de forma autónoma el entorno tecnológico-teórico donde ejercitar las técnicas. Como conclusión, será el propio alumno el encargado de institucionalizar los distintos aspectos del objeto matemático. El profesor sólo intervendrá para aclarar algunos conceptos, si es necesario, o rectificar conceptos erróneos adquiridos.

Metodología para su implementación en el aula

- La resolución de problemas se trabajará en grupos cooperativos de cuatro personas después de hacer una lectura individual del enunciado de cada problema.
- El alumno utilizará la calculadora, así como su cuaderno de clase.
- El profesor supervisará la respuesta dinámica de la clase y atenderá las posibles dudas orientando a los alumnos mientras trabajan.
- Al final de la clase, se hará una puesta en común de las repuestas y se insistirá en aquellos conceptos que supongan alguna dificultad, dejando que sean los propios alumnos los que se aclaren las dudas entre sí mientras el profesor hace de mediador.

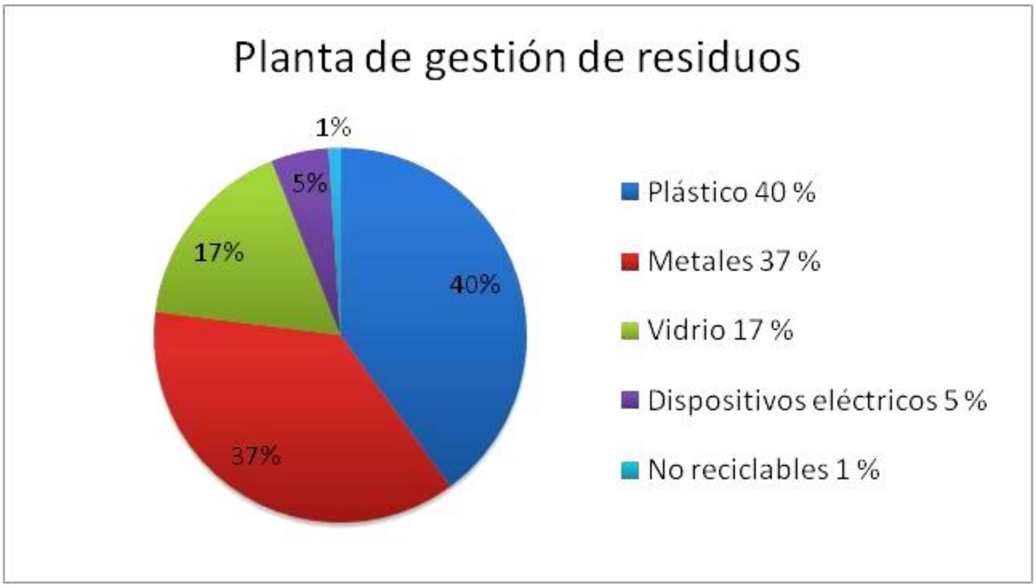
2. **Segundo campo de problemas: organización visual de los datos estadísticos (1 sesión)**

Problema 1. (Oxford Educación – enunciado del problema 11)

Una planta de gestión de residuos se encarga de clasificar y separar los materiales valorizables, esto es, aquellos que pueden reciclarse. Los que no lo son se llevan a aprovechamiento energético o a vertederos. El gráfico muestra el tanto por ciento de los materiales clasificados un día en el que se trataron 4000 toneladas de residuos.

a) Determina la cantidad de cada uno se los diferentes materiales valorizables.

b) Averigua la medida del ángulo que corresponde a cada sector.



Solución:

a) y b)

Plástico 40% $\rightarrow 4000 \cdot 0{,}4 = 1600\ toneladas;\ 0{,}4 \cdot 360° = 144°$

Metales 37% $\rightarrow 4000 \cdot 0{,}37 = 1480\ toneladas;\ 0{,}37 \cdot 360° = 133{,}2°$

Vidrio 17% $\rightarrow 4000 \cdot 0{,}17 = 680\ toneladas;\ 0{,}17 \cdot 360° = 61{,}2°$

Dispositivos eléctricos 5% $\rightarrow 4000 \cdot 0{,}05 = 200\ toneladas;\ 0{,}05 \cdot 360° = 18°$

No reciclables 1% $\rightarrow 4000 \cdot 0{,}01 = 40\ toneladas;\ 0{,}01 \cdot 360° = 3{,}6°$

Problema 2. (Oxford Educación – enunciado del problema 42)

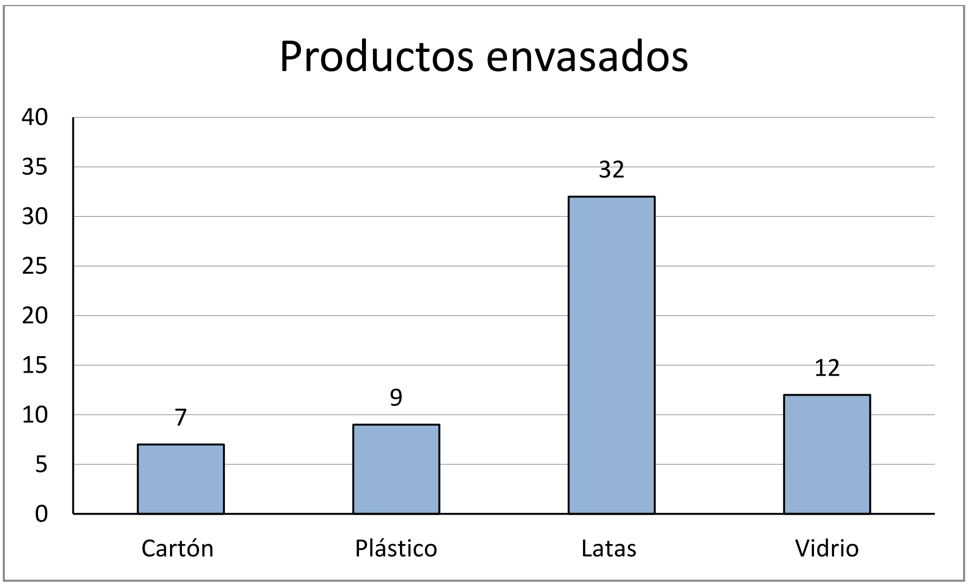
Un cliente ha llenado su carro de la compra con productos envasados. La tabla muestra el tipo de envase y la cantidad de cada uno de ellos.

	Recuento
Cartón	**7**
Plástico	**9**
Latas	**32**
Vidrio	**12**
Total	**60**

Representa los datos de la tabla mediante un diagrama de barras y un diagrama de sectores.

Solución:

Representación de los datos mediante un diagrama de barras.



Representación de los datos mediante un diagrama de sectores.

7 envases de Cartón $\rightarrow (7 \cdot 100)/60 = 11,\hat{6}\%$; $0{,}116 \cdot 360° = 42°$

9 envases de Plástico $\rightarrow (9 \cdot 100)/60 = 15\%$; $0{,}15 \cdot 360° = 54°$

32 envases de Lata $\rightarrow (32 \cdot 100)/60 = 53,\hat{3}\%$; $0{,}533 \cdot 360° = 192°$

12 envases de Vidrio $\rightarrow (12 \cdot 100)/60 = 20\%$; $0{,}2 \cdot 360° = 72°$

<u>Problema 3.</u> (Savia SM – enunciado del problema 5)

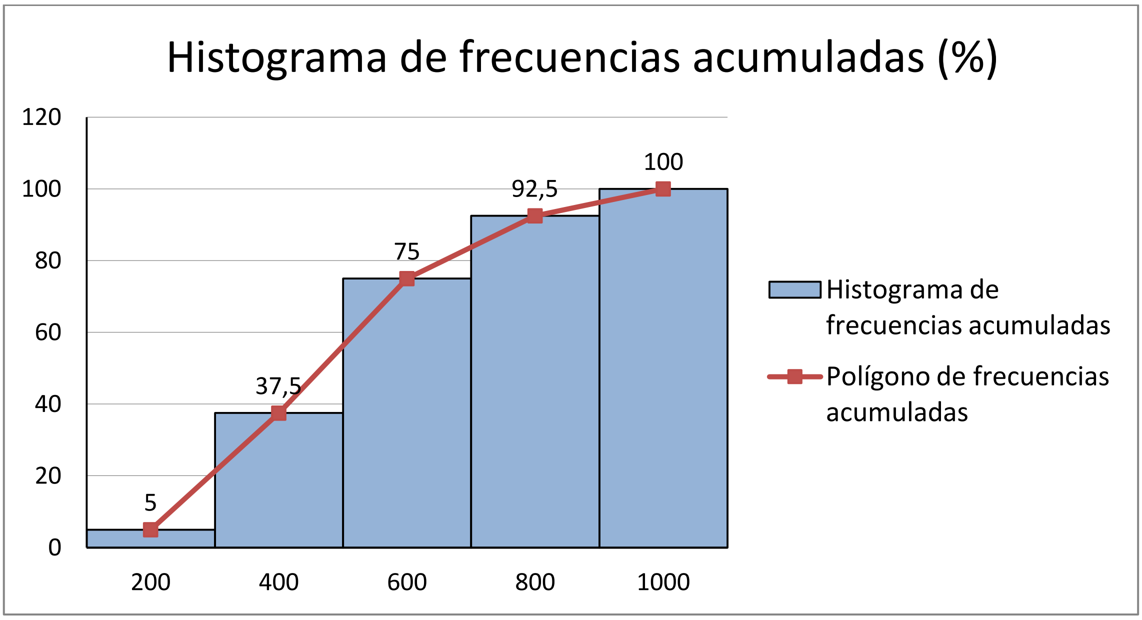
En una fábrica de bombillas se estudia la vida de un tipo de bombilla. Se ha tomado una muestra de 200 lámparas con los siguientes resultados:

Vida en horas	Nº de bombillas
[100,300)	10
[300,500)	65
[500,700)	75
[700,900)	35
[900,1100)	15

Dibuja el polígono de frecuencias acumuladas en porcentaje. ¿Cuál es el porcentaje de bombillas que dura más de 800 horas?

Solución:

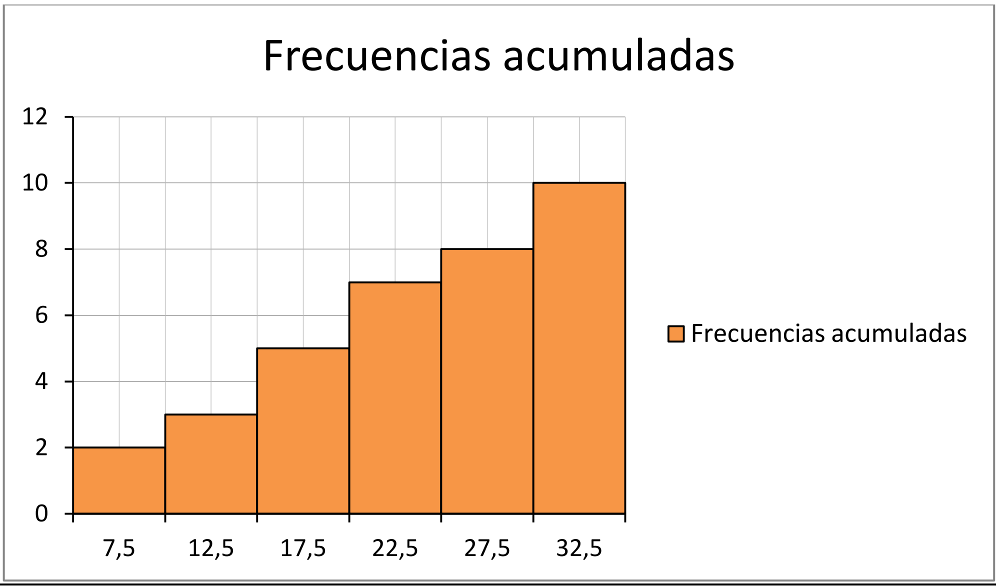
Vida en horas	**Marca de clase**	**Frecuencia absoluta Nº de bombillas**	**Frecuencia relativa**	**Frecuencia relativa acumulada**	**Frecuencia relativa acumulada (%)**
[100,300)	200	**10**	0,05	0,05	5
[300, 500)	400	**65**	0,325	0,375	37,5
[500, 700)	600	**75**	0,375	0,75	75
[700, 900)	800	**35**	0,175	0,925	92,5
[900, 1100)	1000	**15**	0,075	1	100
		200	1		



De la frecuencia relativa acumulada se deduce que % de bombillas que dura más de 800 horas es 100 - 92,5 = 7,5 %.

Problema 4. (Santillana – enunciado del problema 36)

Reconstruye la tabla de frecuencias asociada a este gráfico de frecuencias acumuladas:



Solución:

Dato	Marca de clase	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada	Frecuencia relativa acumulada (%)
[5, 10)	**7,5**	2	0,2	**2**	0,2	20
[10, 15)	**12,5**	1	0,1	**3**	0,3	30
[15, 20)	**17,5**	2	0,2	**5**	0,5	50
[20, 25)	**22,5**	2	0,2	**7**	0,7	70
[25, 30)	**27,5**	1	0,1	**8**	0,8	80
[30, 35)	**32,5**	2	0,2	**10**	1	100
		10				

Problema 5. (Santillana – enunciado del problema 32)

La siguiente tabla muestra los resultados de lanzar 50 veces un dado.

Cara	**1**	**2**	**3**	**4**	**5**	**6**
Nº de veces	**8**	**12**	**5**	**9**	**6**	**10**

a) Representa los diagramas de barras de frecuencias relativas y absolutas. ¿Qué observas?

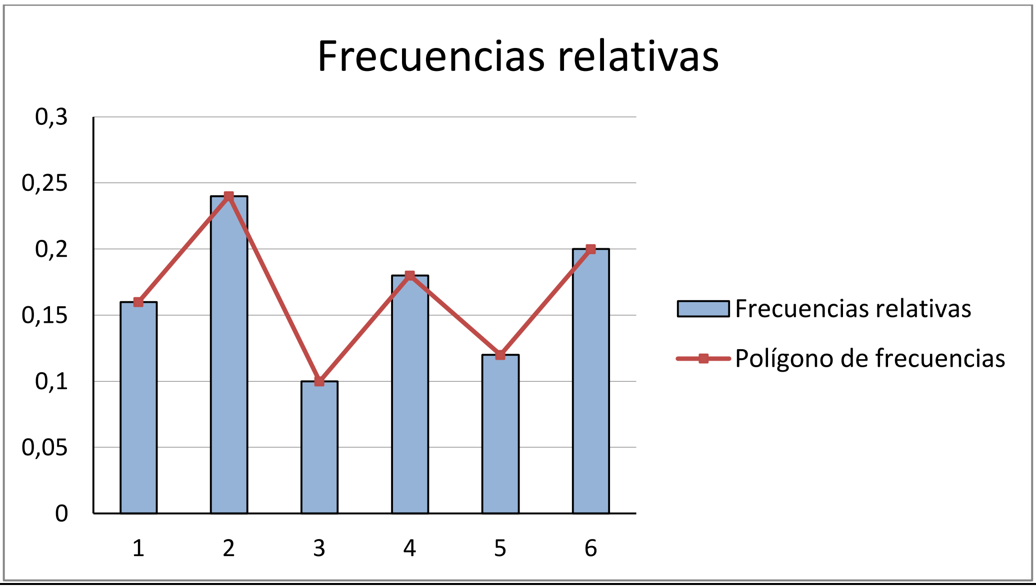
b) Sobre los gráficos anteriores, dibuja su polígono de frecuencias.

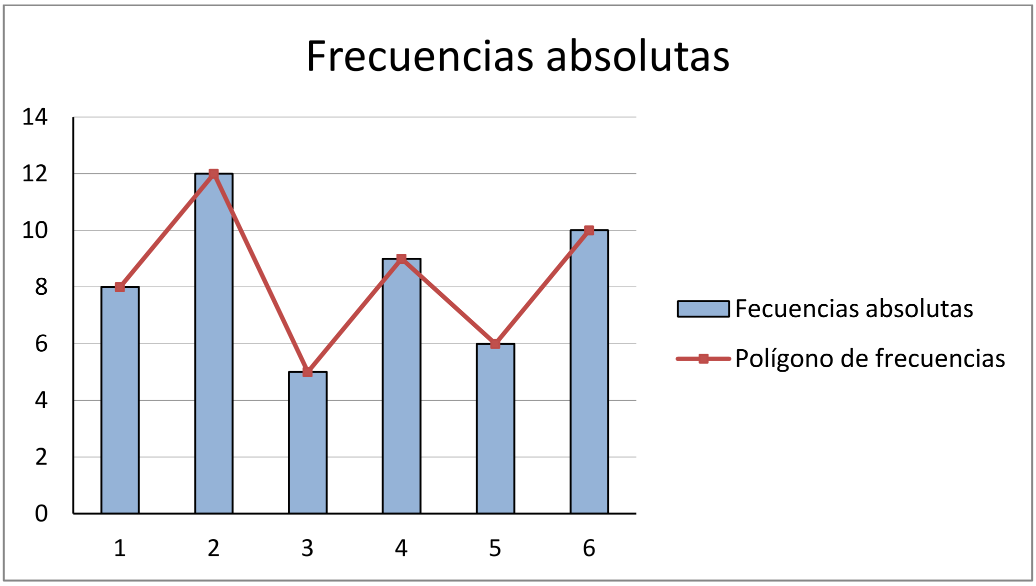
c) ¿Podrías representar los datos en un histograma? Razona tu respuesta.

Solución:

a) y b)

Cara	**1**	**2**	**3**	**4**	**5**	**6**	
Frecuencia absoluta	**8**	**12**	**5**	**9**	**6**	**10**	N = 50
Frecuencia relativa	0,16	0,24	0,1	0,18	0,12	0,2	





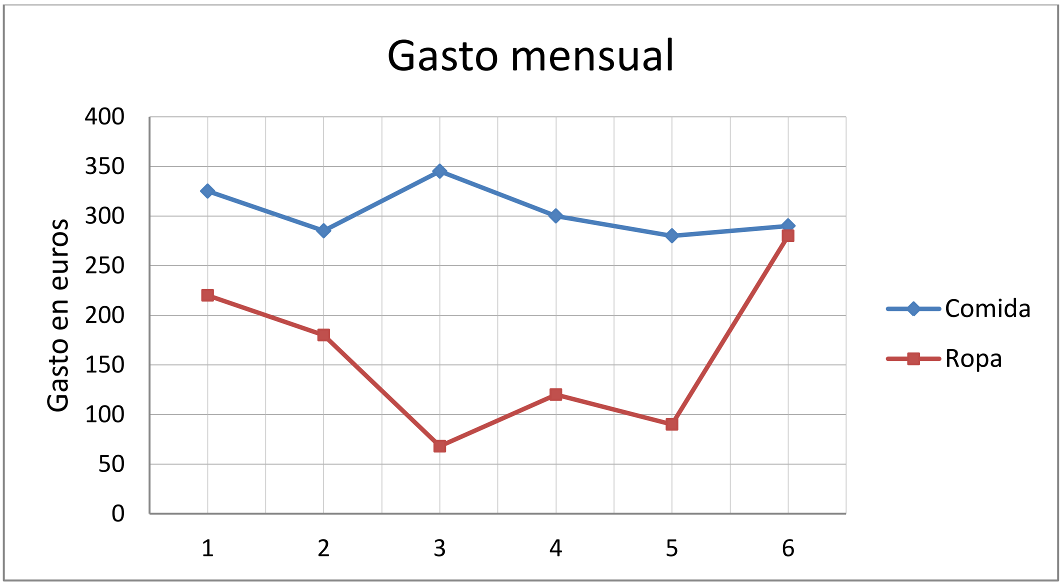
c) No, porque el resultado de lanzar el dado es una variable cuantitativa discreta y los histogramas representan variables cuantitativas continuas.

<u>Problema 6.</u> (Santillana – enunciado del problema 34)

Representa en el gráfico más adecuado el gasto mensual de una familia en comida y en ropa en los últimos 6 meses. ¿Qué conclusiones extraes del gráfico?

Comida	**325**	**285**	**345**	**300**	**280**	**290**
Ropa	**220**	**180**	**68**	**120**	**90**	**280**

<u>Solución:</u>



En el gráfico se observa que el gasto mensual en comida durante los últimos 6 meses se mantiene bastante homogéneo. Sin embargo, el gasto en ropa presenta cambios bruscos de tendencia, con periodos de poco gasto y periodos de mucho gasto que pueden coincidir con los periodos de rebajas o los cambios de estación que se compra más ropa.

Problema 7. (Edelvives – enunciado del problema 3)

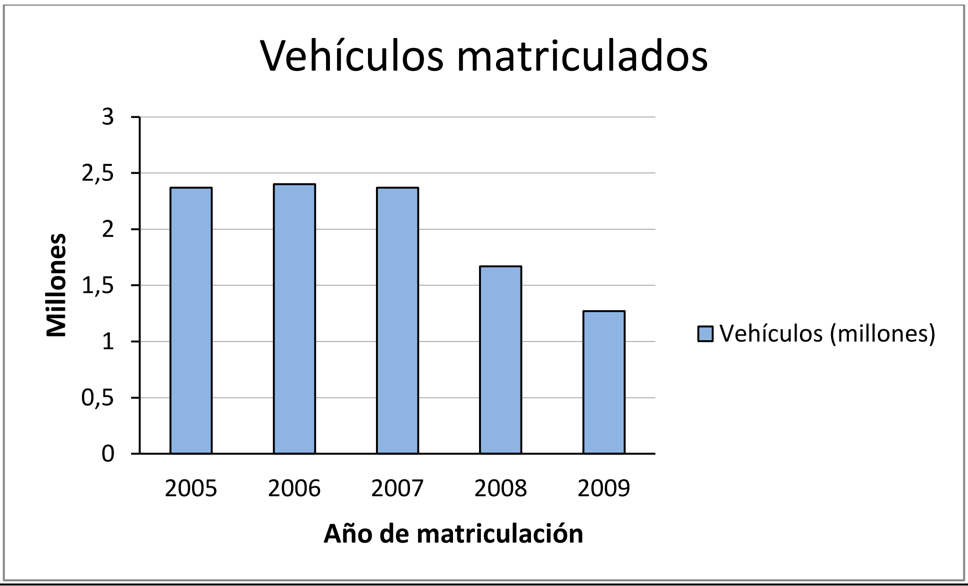
En la siguiente tabla se registran las cifras de automóviles matriculados en cierta comunidad en función del año:

Año	**2005**	**2006**	**2007**	**2008**	**2009**
Vehículos (millones)	**2,37**	**2,40**	**2,37**	**1,67**	**1,27**

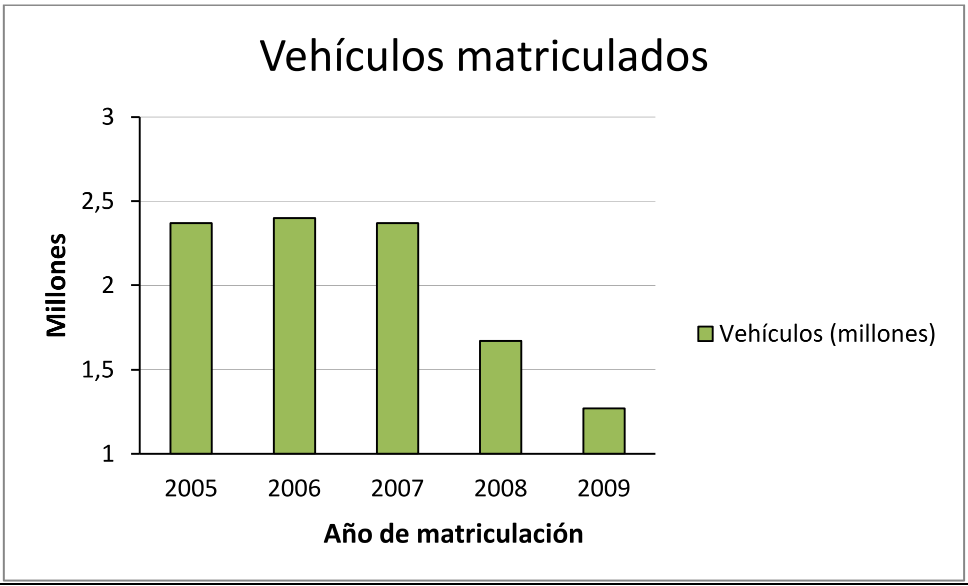
a) Elabora dos diagramas de barras: uno que se ajuste a la regla de los tres cuartos y otro que la incumpla.
b) Describe la diferencia e información que se transmite en una primera impresión.

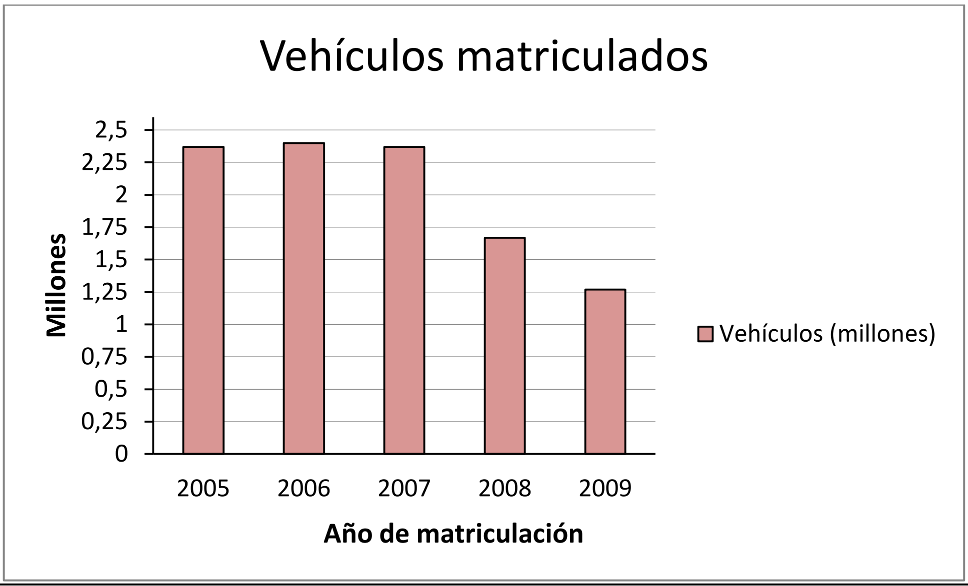
Solución:

a) Diagrama de barras que se ajusta a la regla de los tres cuartos.



A continuación dos ejemplos que incumplen la regla de los tres cuartos.





b) Al incumplir la regla de los tres cuartos, en el primer caso se observa como las diferencias se reducen al comparar los datos, y en el segundo caso las diferencias se magnifican. De este modo, se manipula la interpretación estadística de los gráficos.

Técnicas o modificaciones de una técnica que se ejercitan

En esta sesión se amplían las técnicas anteriores. Se utilizan las tablas de datos, las tablas de frecuencias y los gráficos para representar dichos datos. Se trabajan las siguientes técnicas:

- saber dividir un círculo en sectores de modo que el ángulo de cada sector sea proporcional a la frecuencia correspondiente de cada variable;
- saber representar los datos de modo que la altura de cada barra sea proporcional a la frecuencia correspondiente de cada variable;
- hallar el punto central o marca de clase de un intervalo y representar el polígono de frecuencias asociado;
- la representación gráfica de la evolución temporal de la frecuencia en una variable.

Tecnologías y proceso de institucionalización de los distintos aspectos del objeto matemático

Se amplían las tecnologías de la sesión anterior y se aprende lo siguiente:

- la elección del tipo de gráfico más adecuado para representar una variable cualitativa o cuantitativa (continua o discreta);
- interpretar los gráficos estadísticos;
- reconocer errores y falacias en gráficos estadísticos;
- saber pasar de una representación tabular a una representación gráfica y viceversa.

En esta sesión, a medida que se vayan resolviendo los problemas, el profesor será el que oriente el aprendizaje y asumirá la responsabilidad de institucionalizar los tipos de gráficos adecuados a cada variable.

Se pondrá especial atención al diagrama de sectores, ya que su resolución supone modificaciones de las técnicas iniciales de la primera sesión.

Respecto al reconocimiento de falacias en gráficos estadísticos, el profesor mediará en el debate que pueda surgir y sólo intervendrá para hacer visibles aquellos aspectos que puedan pasar desapercibidos.

Metodología para su implementación en el aula

- Esta sesión se llevará a cabo en el aula de informática y la representación de los gráficos se hará con la hoja de cálculo.
- Se trabajará por parejas en cada ordenador.
- El alumno utilizará la calculadora, así como su cuaderno de clase.
- El profesor dará unas indicaciones previas sobre el uso de gráficos con la hoja de cálculo.
- Aquellos problemas que impliquen un informe de conclusiones, o que puedan dar lugar a distintas respuestas se pondrán en debate. El profesor hará una selección de las aportaciones más significativas y pedirá al alumno que las explique al resto de la clase.

3. Tercer campo de problemas: medidas de centralización y de posición (1 sesión)

Problema 1. (Oxford Educación – enunciado del problema 15)

La media aritmética del número de retrasos de cuatro alumnos a primera hora de la mañana en la última semana ha sido de 3. Ángela llegó tarde 2 días; Pedro, 1 día, y Míriam, los 5 días. ¿Cuántas veces se retrasó Noelia?

Solución:

$$3 = \frac{2+1+5+x}{4} \rightarrow 3 \cdot 4 = 8 + x \rightarrow 12 - 8 = x \rightarrow 4 = x$$

	Dato x_i	Frecuencia absoluta f_i
Ángela	**2**	**1**
Pedro	**1**	**1**
Míriam	**5**	**1**
Noelia	**x**	**1**
		N = 4

<u>Problema 2.</u> (Oxford Educación – enunciado del problema 49)

Calcula la media aritmética de los datos:

1 1 2 2 2 3 3 3 5 6 6 7 7 8 9

a) ¿Cuánto valdrá la media si a cada dato le sumamos 3 unidades?

b) Si cada dato inicial lo multiplicamos por 2, ¿cómo variará la media?

<u>Solución:</u>

Primero calculamos la media aritmética de los datos.

Dato	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	
F. absoluta	2	3	3	1	2	2	1	1	N = 15

$$\bar{x} = \frac{\mathbf{1}\cdot 2 + \mathbf{2}\cdot 3 + \mathbf{3}\cdot 3 + \mathbf{5}\cdot 1 + \mathbf{6}\cdot 2 + \mathbf{7}\cdot 2 + \mathbf{8}\cdot 1 + \mathbf{9}\cdot 1}{15} = 4,\hat{3}$$

a) Se observa que **si sumamos 3 unidades a cada uno de los 15 datos, la media aritmética aumenta en 3 unidades respecto a la inicial.** Es fácil darse cuenta de que al asociar los 15 sumandos "3" y dividir por número total de datos 15, el resultado es sumar 3 unidades a la media aritmética.

$$\bar{x} = \frac{(\mathbf{1}\cdot 2 + \mathbf{2}\cdot 3 + \mathbf{3}\cdot 3 + \mathbf{5}\cdot 1 + \mathbf{6}\cdot 2 + \mathbf{7}\cdot 2 + \mathbf{8}\cdot 1 + \mathbf{9}\cdot 1) + 15\cdot 3}{15} = 4,\hat{3} + 3$$

b) En este caso, **si multiplicamos por 2 cada uno de los datos, la media aritmética resultante es 2 veces la media inicial**. Esto es consecuencia de sacar factor común "2" a todos los sumandos.

$$\bar{x} = \frac{(\mathbf{1}\cdot 2 + \mathbf{2}\cdot 3 + \mathbf{3}\cdot 3 + \mathbf{5}\cdot 1 + \mathbf{6}\cdot 2 + \mathbf{7}\cdot 2 + \mathbf{8}\cdot 1 + \mathbf{9}\cdot 1)\cdot 2}{15} = 4,\hat{3}\cdot 2$$

<u>Problema 3.</u> (Santillana – enunciado del problema 15)

Decide qué valores podemos añadir a este conjunto de datos: 18, 8, 7, 9, 12, 15, 21 y 12 para que la mediana siga siendo la misma.

<u>Solución:</u>

Primero calculamos la mediana del conjunto de datos.

Datos	7	8	9	12	15	18	21	
F. absoluta	1	1	1	2	1	1	1	N = 8
F. absoluta acumulada	1	2	3	5	6	7	8	

La mitad de los datos es N = 8 / 2 = 4.

La primera frecuencia absoluta acumulada mayor o igual que la mitad de los datos es 5 > 4, por tanto la mediana es $M_e = 12$. La mediana es el valor que ocupa la posición central de los datos e indica que hay el mismo número de datos por encima que por debajo de 12.

Debemos añadir datos con la condición de que la mediana siga siendo la misma $M_e = 12$, para ello debemos fijarnos cómo varía el número total de datos, la posición que ocupan los datos y los valores de las frecuencias acumuladas.

<u>Posibilidad 1:</u>

Datos	z < 7	7	8	9	12	15	18	21	
F. absoluta	1	1	1	1	2	1	1	1	N = 9
F. absoluta acumulada	1	2	3	4	6	7	8	9	

Podemos añadir 1 dato con frecuencia 1 por debajo del dato menor 7.

La mitad de los datos es N = 9 / 2 = 4,5.
La primera frecuencia absoluta acumulada mayor o igual que la mitad de los datos es 6; por tanto la mediana es $M_e = 12$.

Si añadimos más datos por debajo del dato menor, la mediana se desplazaría hacía la izquierda.

Posibilidad 2:

Datos	7	8	9	12	15	18	21	x > 21	
F. absoluta	1	1	1	2	1	1	1	1 ó 2	N = 9 ó 10
F. absoluta acumulada	1	2	3	5	6	7	8	9 ó 10	

Podemos añadir 1 dato con frecuencia 1 ó 2, o bien podemos añadir 2 datos con frecuencia 1 cada uno por encima del dato mayor 21.

La mitad de los datos es N = 9 / 2 = 4,5 ó N = 10 / 2 = 5.

La primera frecuencia absoluta acumulada mayor o igual que la mitad de los datos es 5; por tanto la mediana es $M_e = 12$.

Si añadimos más datos por encima del dato mayor, la mediana se desplazaría hacía la derecha.

Problema 4. (Santillana – enunciado del problema 41)

Se ha realizado un estudio entre 200 espectadores para determinar el grado de satisfacción de un programa, obteniendo estos resultados:

Opinión	Muy bueno	Bueno	Regular	Malo	Muy malo
Porcentaje	**15**	**25**	**30**	**25**	**5**

Calcula e interpreta las medidas de centralización.

Solución:

En este problema hay que valorar la opinión en una escala del 1 al 5 para dar una respuesta y darse cuenta de que el porcentaje es la frecuencia relativa en %. Se reconstruye la tabla a partir de la frecuencia relativa y del número total de datos N = 200.

Opinión	Escala del 1 al 5	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada	Frecuencia relativa acumulada (%)
Muy Bueno	5	30	**0,15**	30	0,15	15
Bueno	4	50	**0,25**	80	0,40	40
Regular	3	60	**0,30**	140	0,70	70
Malo	2	50	**0,25**	190	0,95	95
Muy Malo	1	10	**0,05**	200	1	100
		N = 200	1			

Calculamos las medidas de centralización teniendo en cuenta la tabla de frecuencias.

$$Media \rightarrow \bar{x} = \frac{5 \cdot 30 + 4 \cdot 50 + 3 \cdot 60 + 2 \cdot 50 + 10 \cdot 1}{200} = 3{,}2$$

La media aritmética nos indica que la opinión promedio entre los espectadores es regular.

La moda es la opinión que tiene mayor frecuencia, es decir, $M_o = Regular$.

Para calcular la mediana hallamos la mitad de los datos, N = 200 / 2 = 100 y nos fijamos en las frecuencias absolutas acumuladas. La primera frecuencia absoluta acumulada mayor o igual que la mitad de los datos es 140; por tanto, la mediana es $M_e = Regular$.

Problema 5. (Savia SM – enunciado adaptado del problema 8)

La evaluación de un test realizado a los 215 trabajadores de una empresa ha sido:

Nota	1	2	3	4	5	6	7	8	9	10
Nº Trabajadores	9	7	8	34	40	37	50	13	10	7

a) Completa la tabla de frecuencias.

b) Calcula la moda, la media y la mediana.

c) Dibuja el polígono de frecuencias acumuladas.

d) Calcula los cuartiles.

e) Dibuja el diagrama de cajas y bigotes.

f) Deberán hacer un curso de formación los trabajadores por debajo del tercer decil. ¿Cuántos son?

Solución:

a)

Nota	Frecuencia absoluta	Frecuencia relativa	Frecuencia absoluta acumulada	Frecuencia relativa acumulada	Frecuencia relativa acumulada (%)
1	**9**	0,042	9	0,042	4,2
2	**7**	0,032	16	0,074	7,4
3	**8**	0.037	24	0,111	11,1
4	**34**	0,16	58	0,271	27,1
5	**40**	0,186	98	0,457	45,7
6	**37**	0,172	135	0,629	62,9
7	**50**	0,232	185	0,861	86,1
8	**13**	0,060	198	0,921	92,1
9	**10**	0,046	208	0,967	96,7
10	**7**	0,032	215	1	100
	N = 215	1			

b) Calculamos la media $\bar{x}$.

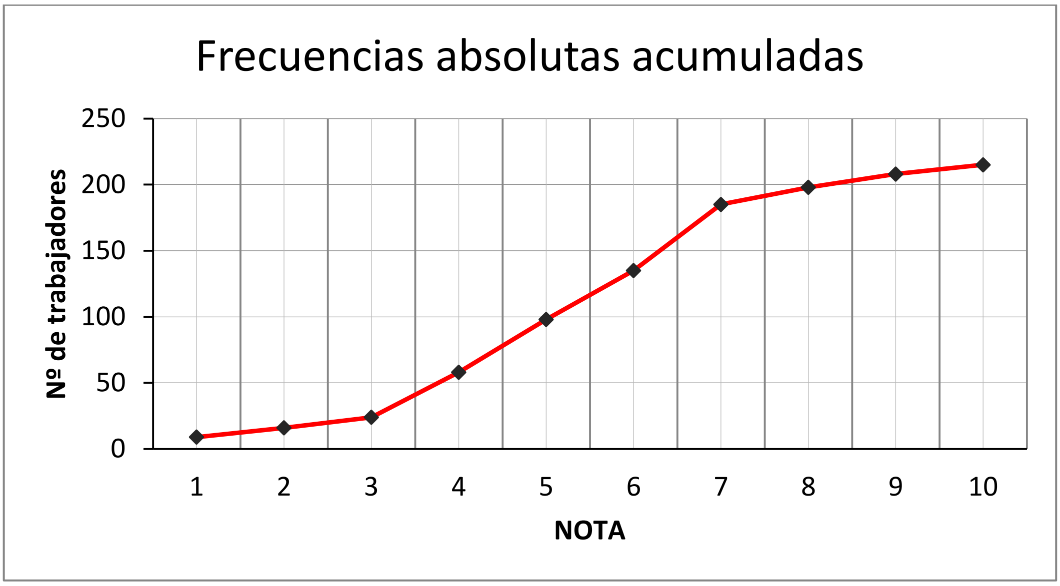
$$\bar{x} = \frac{(\mathbf{1} \cdot 9 + \mathbf{2} \cdot 7 + \mathbf{3} \cdot 8 + \mathbf{4} \cdot 34 + \mathbf{5} \cdot 40 + \mathbf{6} \cdot 37 + \mathbf{7} \cdot 50 + \mathbf{8} \cdot 13 + 9 \cdot 10 + 10 \cdot 7)}{215}$$

$$= \frac{1219}{215} = 5{,}67$$

La moda es la nota que tiene mayor frecuencia, es decir, $M_o = 7$.

Para calcular la mediana hallamos la mitad de los datos, N = 215 / 2 = 107,5 y nos fijamos en las frecuencias absolutas acumuladas. La primera frecuencia absoluta acumulada mayor o igual que la mitad de los datos es 135; por tanto la mediana es $M_e = 6$.

c)



d) Para hallar los cuartiles se dividen los datos ordenados del estudio estadístico en cuatro partes iguales.

$$\frac{215}{4} = 53{,}75$$

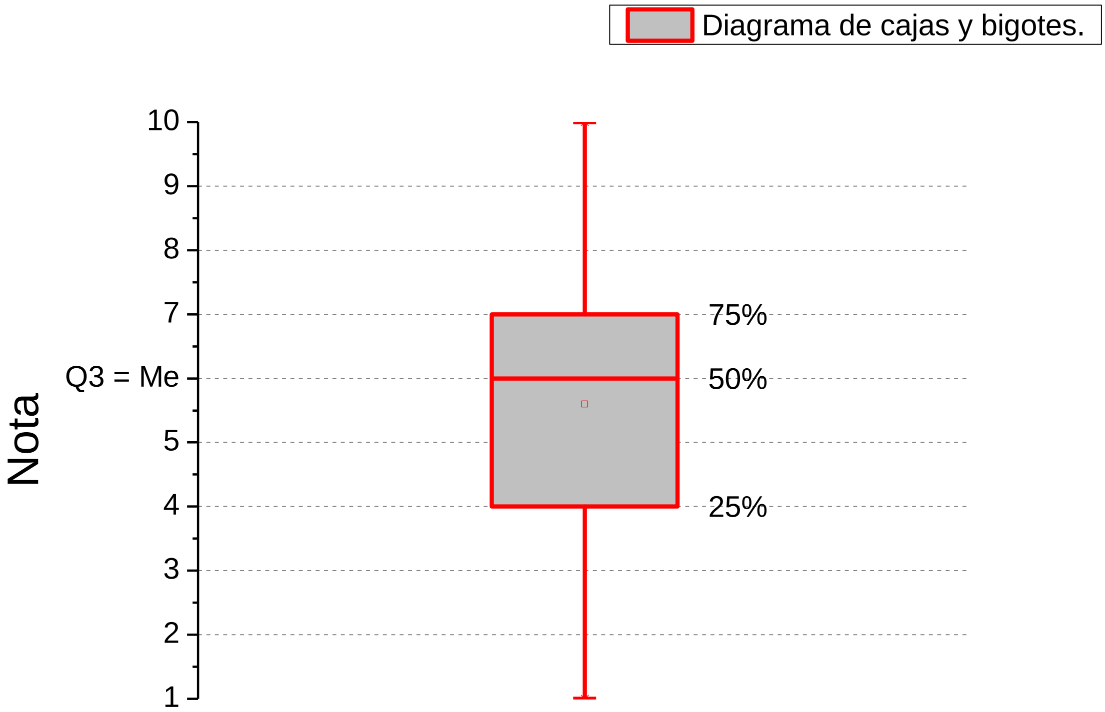
Podemos dividir los 215 datos en 4 partes: 54, 54, 54 y 53.

El primer cuartil, Q_1 , es el dato que es mayor o igual que el 25% de los datos. El 25% de 215 es 53,75. El primer cuartil es el valor de término 54: $Q_1 = x_{54} = 4$.

El segundo cuartil, Q_2 , es el dato que es mayor o igual que el 50% de los datos, es la mediana. El 50% de 215 es 107,5. El segundo cuartil es el valor de término 108: $Q_2 = M_e = x_{108} = 6$.

El tercer cuartil, Q_3 , es el dato que es mayor o igual que el 75% de los datos. El 75% de 215 es 161,25. El tercer cuartil es el valor de término 162: $Q_3 = x_{162} = 7$.

e) Dibujamos el diagrama de cajas y bigotes en función de los cuartiles obtenidos en el apartado d).



f) Si se divide el conjunto de datos ordenados en 10 partes iguales, los valores que separan cada una de las partes se llaman deciles, D_i = 215/10 = 21,5. Por tanto son 65 trabajadores.

Técnicas o modificaciones de una técnica que se ejercitan

En esta sesión se trabaja con tablas de datos y de frecuencias, se introducen los diagramas de bigotes y además, se emplean otras técnicas conocidas ya por los alumnos:

- resolver una ecuación de primer grado;
- asociar sumandos *(propiedad asociativa de la suma)*;
- sacar factor común a una serie de sumandos;
- modificar los datos y estudiar cómo se desplaza la posición de los parámetros estadísticos;
- representar gráficamente los datos mediante un polígono de frecuencias acumuladas;
- representar gráficamente los datos en un diagrama de cajas y bigotes;
- calcular porcentajes.

Tecnologías y proceso de institucionalización de los distintos aspectos del objeto matemático

- Con los problemas 1 y 2 se repasa el concepto de media aritmética. Como es un concepto bastante intuitivo, y no se espera dificultad en su aprendizaje, el propio alumno institucionalizará este concepto.
- La mediana es un concepto aprendido en cursos anteriores, pero en este caso se espera encontrar más dificultad en su aprendizaje. Por ello se plantea el problema 3, para que el alumno sea capaz de aprender por descubrimiento este concepto. Se elegirán las aportaciones más significativas y serán los propios alumnos los que expliquen cómo lo han resuelto.
- En el problema 4 se pretende aprender a interpretar los parámetros estadísticos (media, mediana y moda) en función de los resultados. En este problema se cuantifica una variable cualitativa, como es la opinión, para poder hacer un estudio estadístico sobre el programa. Se espera que la resolución de este estudio estadístico, sobre una encuesta de valoración, sea significativa para el alumno.
- Por último, en el problema 5 se introducen los cuartiles, cuya técnica reside en el cálculo de porcentajes y en encontrar su valor correspondiente en la columna de frecuencias acumuladas. El problema es bastante completo y contiene todos los conceptos vistos hasta ahora. En este proceso será el profesor el que institucionalice los nuevos conceptos.

Metodología para su implementación en el aula

- La resolución de problemas de esta sesión se hará en el aula ordinaria y en grupos colaborativos de cuatro personas.
- Se distribuirá el tiempo para la realización de cada problema según el grado de dificultad o esfuerzo.
- El alumno utilizará la calculadora, así como su cuaderno de clase.
- Tras la realización de cada problema se hará una puesta en común y se dedicará más tiempo a aquellos aspectos más relevantes o complicados para el alumno.

4. Cuarto campo de problemas: medidas de dispersión (1 sesión)

Problema 1. (Savia SM - ejemplos 1 y 2)

a) Calcula el rango, la varianza y la desviación típica de la edad de los profesores de un centro recogidas en la siguiente tabla:

Edad	**[25-35)**	**[35-45)**	**[45-55)**	**[55-65)**
F. absoluta f_i	**6**	**14**	**18**	**12**

b) La media de la edad de los profesores en otro centro es $\bar{x}$ = 38,9 años y la desviación típica es σ = 13,7 años. ¿En qué centro es más representativa la media?

Solución:

a) Se completa la tabla para facilitar los cálculos:

Edad	x_i	**F. absoluta f_i**	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
[25-35)	30	**6**	180	5400
[35-45)	40	**14**	560	22400
[45-55)	50	**18**	900	45000
[55-65)	60	**12**	720	43200
		$N = \Sigma f_i = 50$	$\Sigma x_i \cdot f_i = 2360$	$\Sigma x_i^2 \cdot f_i = 116000$

El recorrido o rango es la diferencia entre el mayor y menor de los valores que toma la variable x_i : R = 65-25 = 40.

Se calcula la media $\bar{x}$ para obtener la varianza σ^2 :

$$\bar{x} = \frac{\Sigma x_i \cdot f_i}{N} = \frac{2360}{50} = 47{,}2 \rightarrow \sigma^2 = \frac{\Sigma x_i^2 \cdot f_i}{N} - \bar{x}^2 = \frac{116000}{50} - 2227{,}84 = 92{,}16$$

Por tanto la desviación típica es:

$$\sigma = \sqrt{\sigma^2} = \sqrt{92{,}16} = 9{,}6 \; años$$

b) Se calculan los coeficientes de variación en los dos centros:

$$CV_1 = \frac{9,6}{47,2} = 0,2033 \rightarrow 20\ \%$$

$$CV_2 = \frac{13,7}{35,9} = 0,3816 \rightarrow 38,16\ \%$$

Como 20% < 38%, la media es más representativa en el primer centro.

<u>**Problema 2.**</u> (Savia SM – enunciado adaptado del problema 10)

El número de libros solicitados por los usuarios de una biblioteca ha sido:

Libros	1	2	3	4	5	6
Usuarios	8	12	9	6	3	2

a) Calcula el recorrido y el recorrido intercuartílico.

b) Halla la varianza y la desviación típica. Interpreta estos parámetros.

<u>Solución:</u>

a) Se completa la tabla para facilitar los cálculos:

Libros x_i	**F. absoluta** f_i	**F. absoluta acumulada** F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
1	**8**	8	8	8
2	**12**	20	24	48
3	**9**	29	27	81
4	**6**	35	24	96
5	**3**	38	15	75
6	**2**	40	12	72
	$N = \Sigma f_i = 40$		$\Sigma x_i \cdot f_i = 110$	$\Sigma x_i^2 \cdot f_i = 380$

El recorrido o rango es la diferencia entre el mayor y menor de los valores que toma la variable x_i : R = 6 - 1 = 5.

Para hallar los cuartiles Q_1 y Q_3 podemos dividir los 40 datos en 4 partes iguales: 10, 10, 10 y 10.

El primer cuartil, Q_1 , es el dato que es mayor o igual que el 25% de los datos. El 25% de 40 es 10. El primer cuartil es el valor de término 10: $Q_1 = x_{10} = 2$.

El tercer cuartil, Q_3 , es el dato que es mayor o igual que el 75% de los datos. El 75% de 40 es 30. El tercer cuartil es el valor de término 30: $Q_3 = x_{30} = 4$.

El recorrido intercuartílico es $R_I = Q_3 - Q_1 = 4 - 2 = 2$.

b) Se calcula la media $\bar{x}$ para obtener la varianza σ^2 :

$$\bar{x} = \frac{\Sigma x_i \cdot f_i}{N} = \frac{110}{40} = 2{,}75 \rightarrow \sigma^2 = \frac{\Sigma x_i^2 \cdot f_i}{N} - \bar{x}^2 = \frac{380}{40} - 7{,}5625 = 1{,}9375$$

Por tanto la desviación típica σ es:

$$\sigma = \sqrt{\sigma^2} = \sqrt{1{,}9375} = 1{,}39\ libros$$

La varianza y la desviación típica nos permiten conocer el grado de agrupamiento de los datos en tono a la media. **Si comparamos la varianza $\boldsymbol{\sigma^2 = 1,9375}$ y la desviación típica**

$\boldsymbol{\sigma = 1,39}$ con la media $\boldsymbol{\bar{x} = 2,75}$, vemos que estos valores son muy grandes lo que implica que los datos están dispersos respecto a las medidas de centralización (media, mediana y moda). En este caso, el número de libros solicitado por un usuario difiere mucho de otro.

<u>Problema 3.</u> (Savia SM – enunciado del problema 37)

Las notas de un grupo de 20 alumnos son:

Nota	**3**	**4**	**5**	**6**	**7**	**8**	**9**
F. absoluta f_i	**2**	**3**	**6**	**4**	**2**	**2**	**1**

a) Halla la media y la desviación típica.

b) Si se supone que se aproxima a una distribución normal, ¿entre qué valores se encuentra aproximadamente el 68% de los datos?

<u>*Solución:*</u>

a) Se completa la tabla para facilitar los cálculos:

Libros x_i	**F. absoluta** f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
3	**2**	6	18
4	**3**	12	48
5	**6**	30	150
6	**4**	24	144
7	**2**	14	98
8	**2**	16	128
9	**1**	9	81
	$N = \Sigma f_i = 20$	$\Sigma x_i \cdot f_i = 111$	$\Sigma x_i^2 \cdot f_i = 667$

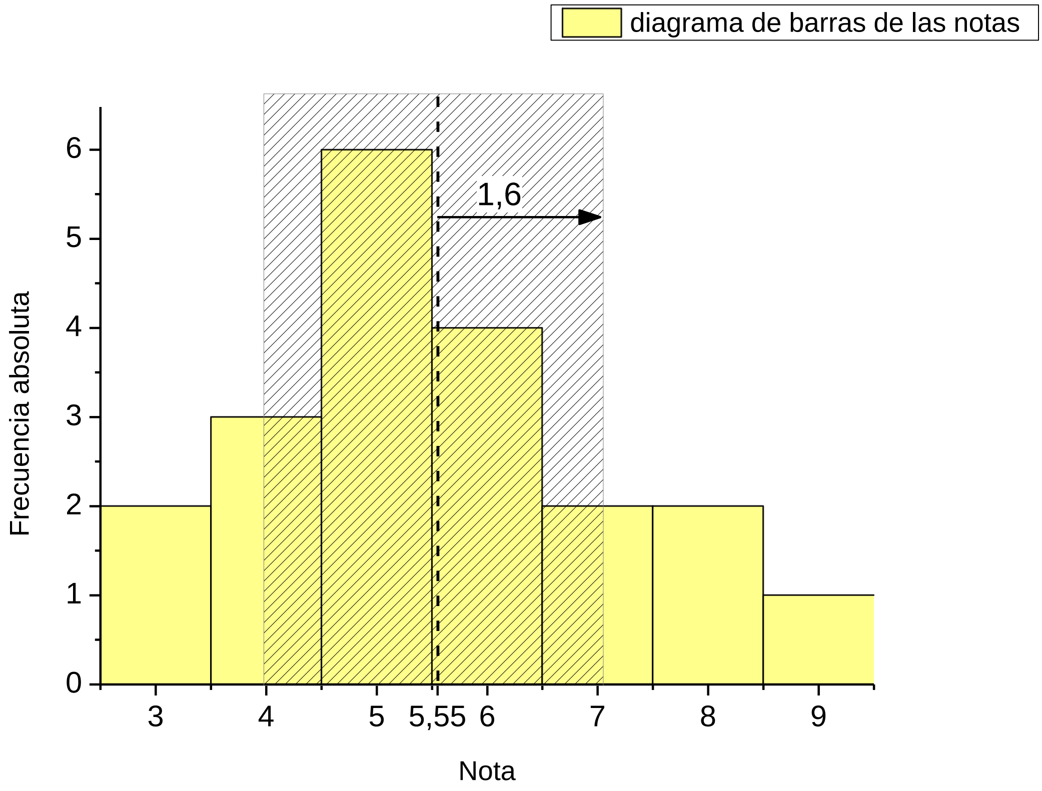
Se calcula la media $\bar{x}$ para obtener la varianza σ^2 :

$$\bar{x} = \frac{\Sigma x_i \cdot f_i}{N} = \frac{111}{20} = 5{,}55 \rightarrow \sigma^2 = \frac{\Sigma x_i^2 \cdot f_i}{N} - \bar{x}^2 = \frac{667}{20} - 30{,}8025 = 2{,}5475$$

Por tanto la desviación típica es:

$$\sigma = \sqrt{\sigma^2} = \sqrt{2{,}5475} = 1{,}5960 \approx 1{,}6$$

b) Se construye el diagrama de barras de las notas.



El 68% de los datos está entre las dos líneas verticales:

$$\bar{x} - \sigma = 5{,}55 - 1{,}6 = 3{,}95$$

$$\bar{x} + \sigma = 5{,}55 + 1{,}6 = 7{,}15$$

Entre 3,95 y 7,15 se encuentran un poco más de las dos terceras partes de las notas.

Problema 4. (Oxford Educación – enunciado adaptado del problema 26)

Las edades de las jugadoras de un equipo de baloncesto son:

15 18 25 20 26 22 26 23 27 25 26 20

a) Calcula e interpreta el coeficiente de variación.

b) Halla el porcentaje de datos que se encuentran en los intervalos:

$$(\bar{x} - \sigma,\ \bar{x} + \sigma), (\bar{x} - 2\sigma,\ \bar{x} + 2\sigma)\ \text{y}\ (\bar{x} - 3\sigma,\ \bar{x} + 3\sigma)$$

Qué conclusiones deduces de los resultados obtenidos.

Solución:

a) Se construye la tabla para facilitar los cálculos:

Edades x_i	**F. absoluta** f_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
15	1	15	225
18	1	18	324
20	2	40	800
22	1	22	484
23	1	23	529
25	2	50	1250
26	3	78	2028
27	1	27	729
	$N = \Sigma f_i = 12$	$\Sigma x_i \cdot f_i = 273$	$\Sigma x_i^2 \cdot f_i = 6369$

Se calcula la media $\bar{x}$ para obtener la varianza σ^2 :

$$\bar{x} = \frac{\Sigma x_i \cdot f_i}{N} = \frac{273}{12} = 22{,}75 \rightarrow \sigma^2 = \frac{\Sigma x_i^2 \cdot f_i}{N} - \bar{x}^2 = \frac{6369}{12} - 517{,}5625 = 13{,}1875$$

Por tanto la desviación típica es:

$$\sigma = \sqrt{\sigma^2} = \sqrt{13{,}1875} = 3{,}631$$

El coeficiente de variación CV es el cociente entre la desviación típica y la media:

$$CV = \frac{\sigma}{\bar{x}} = \frac{3{,}631}{22{,}75} = 0{,}1596$$

El coeficiente de variación representa el número de veces que la desviación típica contiene a la media. Así, cuanto menor sea su valor, menor será la dispersión y mayor la representatividad de la media. Tenemos un 15% de coeficiente de

variación por lo que los datos están agrupados en torno a la media y la media es bastante representativa.

b) El porcentaje de datos que se encuentran en el intervalo $(\mathbf{\bar{x}-\sigma},\ \mathbf{\bar{x}+\sigma})$**:**

$$\bar{x}-\sigma = 22{,}75 - 3{,}631 = 19{,}119$$

$$\bar{x}+\sigma = 22{,}75 + 3{,}631 = 26{,}381$$

En el intervalo $(19{,}119,\ 26{,}381)$ se encuentran los datos 20, 22, 23, 25 y 26, tenemos 9 datos en total que son el 75% de los datos.

$$\%\,datos = \frac{9\cdot 100}{12} = 75\%$$

El porcentaje de datos que se encuentran en el intervalo $(\mathbf{\bar{x}-2\sigma},\ \mathbf{\bar{x}+2\sigma})$**:**

$$\bar{x}-2\sigma = 22{,}75 - 7{,}262 = 15{,}488$$

$$\bar{x}+2\sigma = 22{,}75 + 7{,}262 = 30{,}012$$

En el intervalo $(15{,}488,\ 30{,}012)$ se encuentran los datos 18, 20, 22, 23, 25, 26 y 27, tenemos 11 datos en total que son el $91,\widehat{6}\%$ de los datos.

$$\%\,datos = \frac{11\cdot 100}{12} = 91,\widehat{6}\%$$

El porcentaje de datos que se encuentran en el intervalo $(\mathbf{\bar{x}-3\sigma},\ \mathbf{\bar{x}+3\sigma})$:

$$\bar{x}-3\sigma = 22{,}75 - 10{,}893 = 11{,}857$$

$$\bar{x}+3\sigma = 22{,}75 + 10{,}893 = 33{,}643$$

En el intervalo $(11{,}857,\ 33{,}643)$ se encuentran los datos 15, 18, 20, 22, 23, 25, 26 y 27, tenemos 12 datos en total que son el 100% de los datos.

$$\%\,datos = \frac{12\cdot 100}{12} = 100\%$$

Se puede resumir toda la información de cualquier distribución estadística con solo dos números: la media y la desviación típica. La media indica el centro y la desviación típica el grado de dispersión de los datos. En el problema el 75 % de los

datos están entre $\bar{x} - \sigma$ y $\bar{x} + \sigma$, es decir, casi todos los datos de la distribución. Si nos alejamos más del centro solo aparecen 2 o 3 datos más.

Problema 5. (Santillana – enunciado del problema 52)

Los salarios, en euros, en una empresa son los siguientes:

Mujeres: 1200, 1300, 1000, 900, 900, 1100, 1200, 1100, 1400, 1200, 1000, 1300, 1200, 1100, 1100.

Hombres: 1200, 1300, 1500, 1300, 1400, 900, 1700, 1600, 1400, 1300, 1500, 1300, 1900, 1700, 1200.

a) Calcula la distribución de frecuencias, la media, la mediana y la moda, de cada grupo: hombres y mujeres.

b) Calcula sus medidas de dispersión: rango o recorrido, varianza, desviación típica y coeficiente de variación.

c) Compara ambos grupos. ¿Cómo lo haces?

d) Si consideramos todos los datos en el mismo grupo, ¿qué resultados obtenemos?

Solución:

a.1) Se construye *la tabla del grupo de mujeres* para facilitar los cálculos:

Salario M. x_i	**F. absoluta** f_i	**F. absoluta acumulada**	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
900	2	2	1800	1620000
1000	2	4	2000	2000000
1100	4	8	4400	4840000
1200	4	12	4800	5760000
1300	2	14	2600	3380000
1400	1	15	1400	1960000
	$N = \Sigma f_i = 15$		$\Sigma x_i \cdot f_i = 17000$	$\Sigma x_i^2 \cdot f_i$ =19560000

Calculamos la media $\bar{x}$:

$$\bar{x} = \frac{\Sigma x_i \cdot f_i}{N} = \frac{17000}{15} = 1133,\hat{3}$$

La moda es el salario que tiene mayor frecuencia. Puede haber más de una moda, en este caso, $M_{o1} = 1100 \; y \; M_{o2} = 1200$.

Para calcular la mediana hallamos la mitad de los datos, N = 15 / 2 = 7,5 y nos fijamos en las frecuencias absolutas acumuladas. La primera frecuencia absoluta acumulada mayor o igual que la mitad de los datos es 8, por tanto, la mediana es $M_e = 1100$.

b.1) Se calcula la varianza σ^2 *del grupo de mujeres*:

$$\sigma^2 = \frac{\Sigma x_i^2 \cdot f_i}{N} - \bar{x}^2 = \frac{19560000}{15} - 1284444,\hat{4} = 19555,6$$

Por tanto la desviación típica es:

$$\sigma = \sqrt{\sigma^2} = \sqrt{19555,6} = 139,84$$

El coeficiente de variación CV es el cociente entre la desviación típica y la media:

$$CV = \frac{\sigma}{\bar{x}} = \frac{139,84}{1133,\hat{3}} = 0,123 \rightarrow 12,3\%$$

a.2) Se construye *la tabla del grupo de hombres* para facilitar los cálculos:

Salario H. x_i	**F. absoluta** f_i	**F. absoluta acumulada**	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
900	1	1	900	810000
1200	2	3	2400	2880000
1300	4	7	5200	6760000
1400	2	9	2800	3920000
1500	2	11	3000	4500000
1600	1	12	1600	2560000
1700	2	14	3400	5780000
1900	1	15	1900	3610000
	$N = \Sigma f_i = 15$		$\Sigma x_i \cdot f_i$ =21200	$\Sigma x_i^2 \cdot f_i = 30820000$

Calculamos la media $\bar{x}$:

$$\bar{x} = \frac{\Sigma x_i \cdot f_i}{N} = \frac{21200}{15} = 1413,\hat{3}$$

La moda es el salario que tiene mayor frecuencia, es decir, $M_o = 1300$.

Para calcular la mediana hallamos la mitad de los datos, N = 15 / 2 = 7,5 y nos fijamos en las frecuencias absolutas acumuladas. La primera frecuencia absoluta acumulada mayor o igual que la mitad de los datos es 9; por tanto la mediana es $M_e = 1400$.

b.2) Se calcula la varianza σ^2 *del grupo de hombres*:

$$\sigma^2 = \frac{\Sigma x_i^2 \cdot f_i}{N} - \bar{x}^2 = \frac{30820000}{15} - 1997511{,}11 = 57249{,}77$$

Por tanto la desviación típica es:

$$\sigma = \sqrt{\sigma^2} = \sqrt{57249{,}77} = 239{,}269$$

El coeficiente de variación CV es el cociente entre la desviación típica y la media:

$$CV = \frac{\sigma}{\bar{x}} = \frac{239,269}{1413,\hat{3}} = 0,169 \rightarrow 16,9\%$$

c) Las medidas de centralización y de dispersión proporcionan mayor información cuando se analizan conjuntamente. Nos fijamos en el coeficiente de variación que relaciona la media de los datos con la desviación típica. **El coeficiente de variación para el grupo de mujeres es del 12,6% y el coeficiente de variación para los hombres es del 16,9%. El coeficiente de variación es pequeño lo que indica que los datos presentan una agrupación bastante grande respecto de la media, la mediana y la moda. En el grupo de mujeres esta agrupación es ligeramente mayor que en el grupo de hombres**

d) Se construye *la tabla de todos los datos en el mismo grupo:*

Salario M y H. x_i	**F. absoluta f_i**	**F. absoluta acumulada**	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
900	3	3	2700	2430000
1000	2	5	2000	2000000
1100	4	9	4400	4840000
1200	6	15	7200	8640000
1300	6	21	7800	10140000
1400	3	24	4200	5880000
1500	2	26	3000	4500000
1600	1	27	1600	2560000
1700	2	29	3400	5780000
1900	1	30	1900	3610000
	$N = \Sigma f_i = 30$		$\Sigma x_i \cdot f_i = 38200$	$\Sigma x_i^2 \cdot f_i = 50380000$

Calculamos la media $\bar{x}$:

$$\bar{x} = \frac{\Sigma x_i \cdot f_i}{N} = \frac{38200}{30} = 1273,\hat{3}$$

La moda es el salario que tiene mayor frecuencia. Puede haber más de una moda, en este caso, $M_{o1} = 1200\ y\ M_{o2} = 1300.$

Para calcular la mediana hallamos la mitad de los datos, N = 30 / 2 = 15 y nos fijamos en las frecuencias absolutas acumuladas. La primera frecuencia absoluta acumulada mayor o igual que la mitad de los datos es 15; por tanto la mediana es $M_e = 1200.$

Se calcula la varianza σ^2 :

$$\sigma^2 = \frac{\Sigma x_i^2 \cdot f_i}{N} - \bar{x}^2 = \frac{50380000}{30} - 1621292{,}89 = 58040{,}44$$

Por tanto la desviación típica es:

$$\sigma = \sqrt{\sigma^2} = \sqrt{58040{,}44} = 240{,}91$$

El coeficiente de variación CV es el cociente entre la desviación típica y la media:

$$CV = \frac{\sigma}{\bar{x}} = \frac{240{,}91}{1273,\hat{3}} = 0{,}189 \rightarrow 18{,}9\%$$

Nos fijamos en el coeficiente de variación que relaciona la media de los datos con la desviación típica. El coeficiente de variación si consideramos todos los datos en el mismo grupo es del 18,9%. **El coeficiente de variación es moderadamente pequeño lo que indica que los datos presentan una agrupación moderadamente grande respecto de la media, la mediana y la moda. Al tratar los datos conjuntamente, la agrupación es menor que si tratamos los datos de mujeres y hombres por separado.**

Técnicas o modificaciones de una técnica que se ejercitan

Durante la sesión se ejercitan todas las técnicas utilizadas anteriormente por lo que nos centraremos en:

- ejercitar y perfeccionar las técnicas aprendidas;
- manejar el lenguaje estadístico y asociar a cada símbolo su significado.

Tecnologías y proceso de institucionalización de los distintos aspectos del objeto matemático

- El aprendizaje de los parámetros de dispersión justifica todas las técnicas que se han ido trabajando hasta ahora. Por esta razón, se espera que el alumno posea cierta soltura en la organización de los datos y los cálculos.
- El profesor institucionalizará los conceptos de recorrido, varianza, desviación típica y coeficiente de variación al resolver los problemas 1 y 2.
- El alumno será el responsable de institucionalizar el concepto de distribución normal a través de la resolución de los problemas 3 y 4, ya que posee los conocimientos necesarios para ello.
- El problema 5 es el más representativo de la sesión. Supone un refuerzo para institucionalizar de nuevo los parámetros de centralización y de dispersión. En este problema se pondrán en debate los resultados para dar una interpretación conjunta de estos parámetros.

Metodología para su implementación en el aula

- Al comienzo de clase se hará un breve repaso de los parámetros de dispersión.
- La resolución de problemas se hará en el aula ordinaria y en grupos colaborativos de cuatro personas.
- Se distribuirá el tiempo para la realización de los problemas.
- Previsiblemente, no se puedan completar todos los problemas en esta sesión debido a su larga extensión, pudiéndose omitir el problema 4 y mandarlo como tarea específica para casa.
- El alumno utilizará la calculadora, así como su cuaderno de clase.
- Al acabar cada problema se hará una puesta en común

- A lo largo de la sesión se insistirá en la interpretación conjunta de la media y la desviación típica.
- Asimismo uno de los objetivos es perfeccionar las técnicas para manejar los datos en formatos tabulares.

5. Quinto campo de problemas: estadística bidimensional (3 sesiones)

Problema 1. (Oxford Educación – enunciado del problema 28)

Fíjate en estos estudios estadísticos:

I. En los hogares, el coste en el recibo de la luz y el número de personas que residen en una casa.

II. En una ciudad, el número de días soleados y la cantidad de agua de lluvia registrada, en L/m^2.

III. En familias con hijos mayores de edad, la estatura de los padres y la estatura de los hijos.

Determina:

a) Cuáles son las variables que se relacionan.

b) El colectivo de individuos que se estudia.

c) El tipo de relación entre las variables.

Solución I: Las variables son el coste del recibo de luz y el número de personas que viven en la misma cas. El colectivo de individuos puede ser los hogares de una ciudad. La relación entre variables es que a mayor número de personas en un mismo hogar, mayor gasto de luz.

Solución II: Las variables son el número de días con sol y la cantidad de agua registrada. El colectivo puede ser los días de un mes o un año en una ciudad determinada. La relación entre variables es que a mayor número de días soleados, menor cantidad de agua registrada.

Solución III: Las variables son la estatura de los padres y la estatura de los hijos. El colectivo puede ser las familias de una ciudad o país. La relación entre variables es que a mayor estatura de los padres, mayor estatura de los hijos.

Problema 2. (Edelvives – enunciado del problema 13)

Se ha preguntado a diversos matrimonios por el número de años que llevan casados y el número de hijos que han tenido desde que contrajeron matrimonio:

Años	**3**	**8**	**6**	**12**	**25**	**13**	**9**	**7**	**16**
Hijos	**1**	**3**	**2**	**0**	**2**	**4**	**1**	**2**	**3**

a) Elabora una tabla simple de los datos.

b) Representa la nube de puntos.

c) ¿Cuántos matrimonios han tenido menos de 2 hijos?

d) ¿Cuántos años llevan casados los matrimonios que han tenido tres hijos?

e) ¿Qué matrimonios han tenido mayor número de hijos: los que llevan casados más de 10 años o los que llevan menos de 10 años?

Solución:

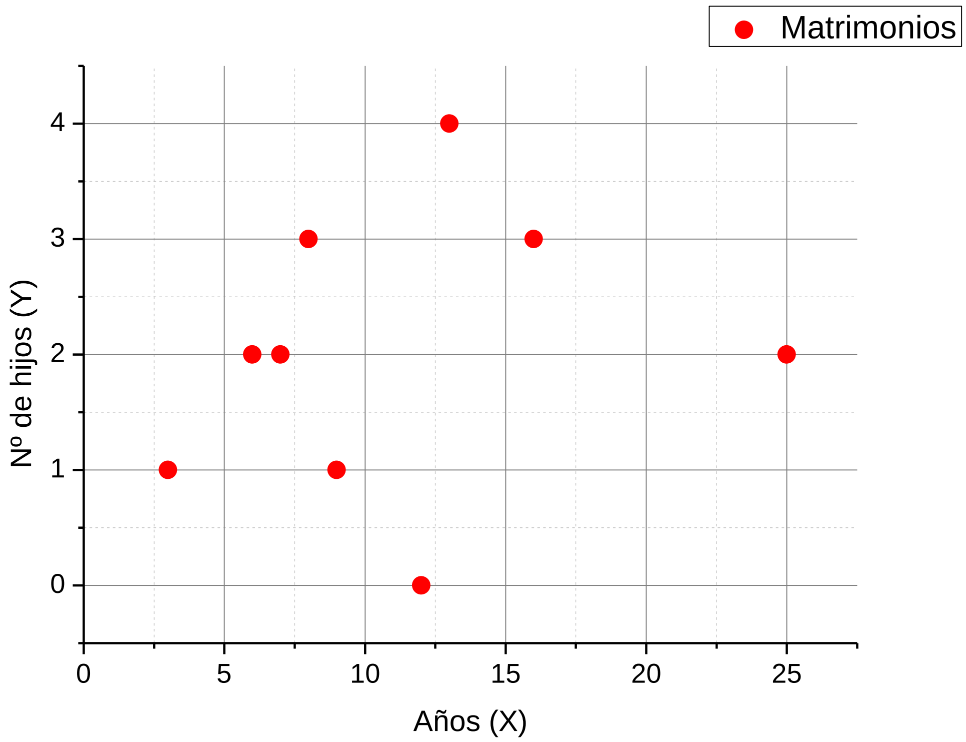
a)

Años (X)	3	6	7	8	9	12	13	16	25
F. absoluta	1	1	1	1	1	1	1	1	1

Nº de hijos (Y)	0	1	2	3	4
F. absoluta	1	2	3	2	1

b) Reconstruimos la tabla del enunciado de forma ordenada para representar la nube de puntos:

Años (X)	3	6	7	8	9	12	13	16	25
Nº de hijos (Y)	1	2	2	3	1	0	4	3	2



c) Nos fijamos en la tabla del enunciado: sumamos los matrimonios que han tenido 0 ó 1 hijo, en total son 3 matrimonios.

d) Nos fijamos en la tabla del enunciado: los años que llevan casados los matrimonios que han tenido tres hijos son 8 y 16 años.

e) **Nos fijamos en la tabla ordenada del apartado b): sumamos el número de hijos de los 4 matrimonios que llevan casados más de 10 años que son 9 y sumamos el número de hijos de los 5 matrimonios que llevan casados menos de 10 años, que son 9 también. Con estos resultados, observamos que los matrimonios que llevan casados más de 10 años tienen mayor número de hijos, se puede comprobar haciendo las medias respectivas.**

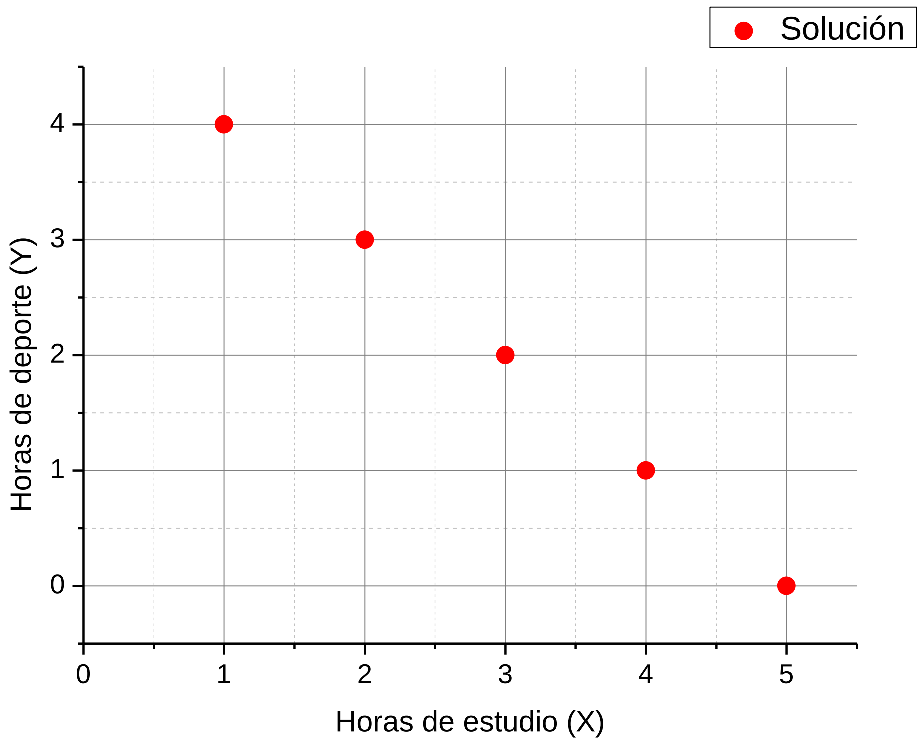
Problema 3. (Oxford Educación – ejemplo1 adaptado)

Representa la nube de puntos correspondiente a la siguiente variable estadística bidimensional y explica qué tipo de relación observas entre las variables que la forman.

Horas de estudio (X)	**1**	**2**	**3**	**4**	**5**
Horas de deporte (Y)	**4**	**3**	**2**	**1**	**0**

Solución:

Se observa una recta de pendiente negativa que se ajusta a todos los puntos. **La relación entre las variables es inversa y funcional, es decir, cuantas más horas se dedican al estudio, menos horas se dedican al deporte.**



Problema 4. (Savia SM – enunciado del problema 16)

Las puntuaciones obtenidas por 40 personas en dos test que miden la comprensión lectora (X) y el cálculo numérico (Y) han sido:

X \ Y	20	30	40	50
20	3	4	2	1
30	5	2	2	1
40	1	3	4	2
50	0	1	3	6

a) ¿Cuántas personas han obtenido 30 puntos en comprensión lectora? ¿Y 50 puntos en cálculo numérico?

b) ¿Hay correlación entre ambas habilidades?

Solución:

a) En comprensión lectora (X) han obtenido 30 puntos 10 personas (= 5+2+2+1) y en cálculo numérico (Y) han obtenido 50 puntos 10 personas (= 1+1+2+6).

b) **Los puntos no se distribuyen en torno a una línea recta, esto significa que no hay correlación entre ambas habilidades: no hay relación entre la comprensión lectora y la habilidad para el cálculo**.

Problema 5. (Oxford Educación – enunciado del problema 32)

Construye la tabla de doble entrada que relacione las horas de estudio (X) con las horas de sueño (Y), de 15 alumnos, si los datos obtenidos en una encuesta han sido:

(1, 7)	**(1, 7)**	**(1, 8)**	**(1, 9)**	**(2, 7)**
(2, 7)	**(2, 7)**	**(2, 7)**	**(2, 8)**	**(2, 9)**
(3, 7)	**(3, 7)**	**(3, 7)**	**(3, 8)**	**(3, 9)**

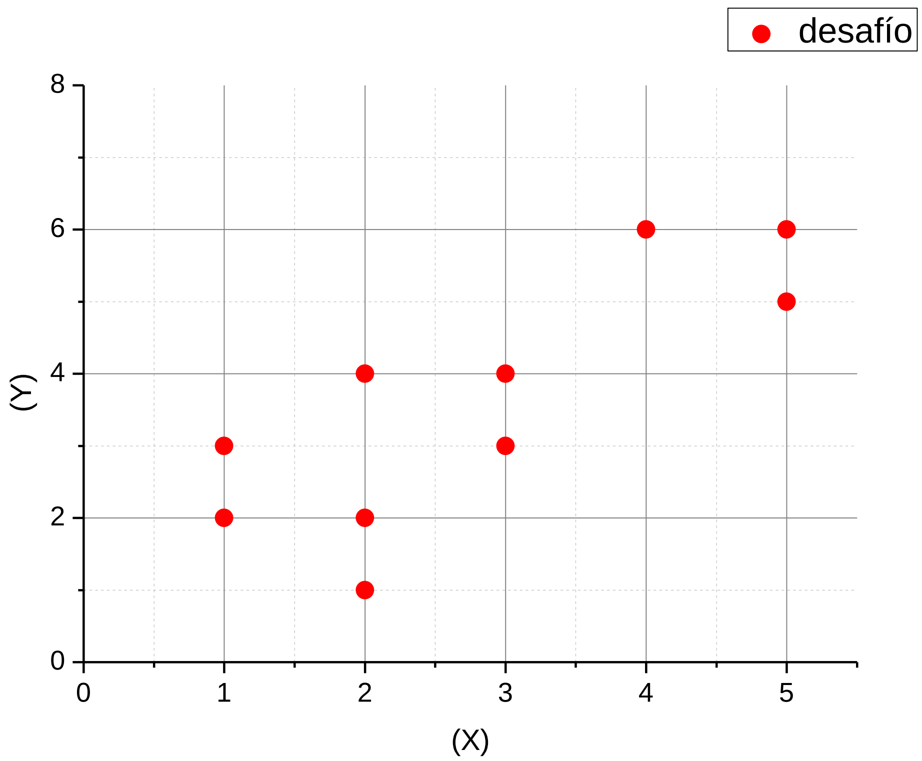
Solución:

La tabla de doble entrada es la siguiente:

X \ Y	7	8	9
1	**2**	**1**	**1**
2	**4**	**1**	**1**
3	**3**	**1**	**1**

Problema 6. (Oxford Educación – enunciado del problema 33)

A la vista del diagrama de dispersión, crea la tabla de doble entrada y las tablas de frecuencias absolutas de las variables X e Y.



Solución:

La tabla de doble entrada es la siguiente:

X	1	1	2	2	2	3	3	4	5	5
Y	2	3	1	2	4	3	4	6	5	6

Las tablas de frecuencias absolutas de las variables X e Y son las siguientes:

X	1	2	3	4	5
F. absoluta	2	3	2	1	2

Y	2	3	1	4	5	6
F. absoluta	2	2	1	2	1	2

Problema 7. (Oxford Educación – enunciado adaptado del problema 36)

La tabla muestra la estatura, en cm, y la talla de calzado de cinco personas.

Estatura (X)	**160**	**165**	**168**	**170**	**175**
Talla de calzado (Y)	**38**	**39**	**40**	**42**	**45**

a) Representa el diagrama de dispersión e indica qué tipo de relación existe entre las variables.

b) Copia la tabla en tu cuaderno y complétala.

x_i	y_i	f_i	$x_i f_i$	$x_i^2 f_i$	$y_i f_i$	$y_i^2 f_i$	$x_i y_i f_i$
160	**38**	**1**					
165	**39**	**1**					
168	**40**	**1**					
170	**42**	**1**					
175	**45**	**1**					
838	**204**						**34250**

c) Ayúdate de la tabla para calcular la media $\bar{x}$ y la media $\bar{y}$, y representa en el diagrama de dispersión el punto $(\bar{x}, \bar{y})$.

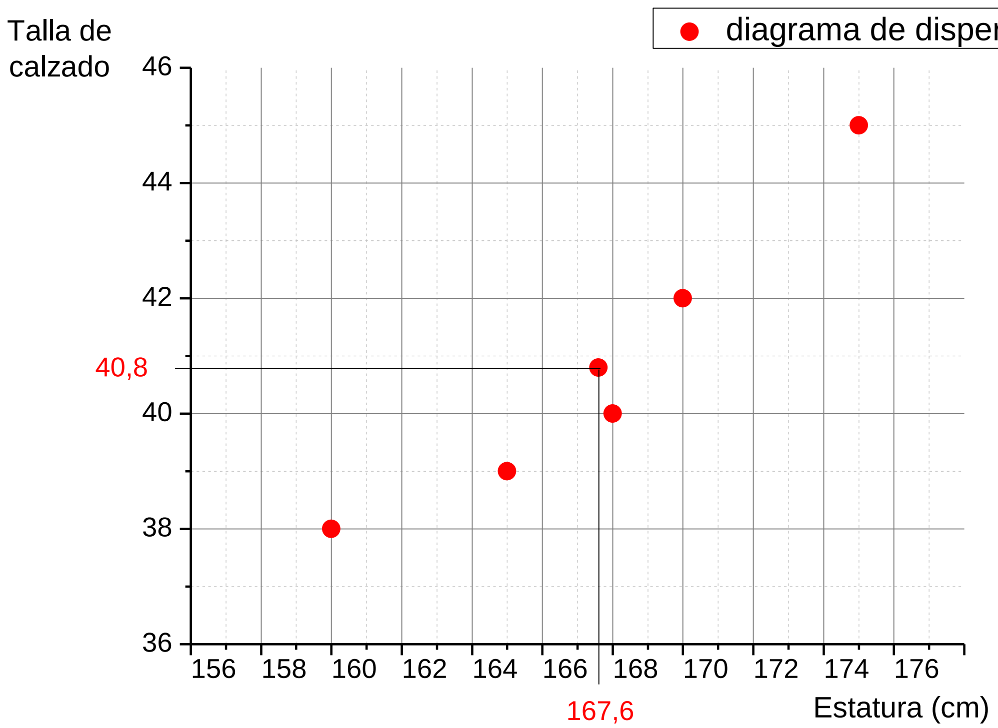
d) Ayúdate de la tabla para calcular la varianza en x y la varianza en y.

e) Ayúdate de la tabla para calcular la covarianza en x e y.

f) Calcula e interpreta el coeficiente de correlación r.

Solución:

a) En el diagrama de dispersión se observa que la relación entre variables se aproxima a una recta de pendiente positiva por lo que el valor del coeficiente de correlación lineal será $0 < r < 1$.



b) Completamos la tabla:

x_i	y_i	f_i	$x_i f_i$	$x_i^2 f_i$	$y_i f_i$	$y_i^2 f_i$	$x_i y_i f_i$
160	**38**	**1**	160	25600	38	1444	6080
165	**39**	**1**	165	27225	39	1521	6435
168	**40**	**1**	168	28224	40	1600	6720
170	**42**	**1**	170	28900	42	1764	7140
175	**45**	**1**	175	30625	45	2025	7875
838	**204**	5	838	140574	204	8354	**34250**

c) Con ayuda de la tabla se calcula la media $\bar{x}$ y la media $\bar{y}$:

$$\bar{x} = \frac{838}{5} = 167{,}6$$

$$\bar{y} = \frac{204}{5} = 40{,}8$$

Se representa en el diagrama de dispersión el punto $(\bar{x}\,, \bar{y}) = (167{,}6\,, 40{,}8)$.

d) Con ayuda de la tabla se calcula la varianza en x y la varianza en y.

$$\sigma_x^2 = \frac{140574}{5} - (167{,}6)^2 = 25{,}04$$

$$\sigma_y^2 = \frac{8354}{5} - (40{,}8)^2 = 6{,}16$$

e) Con ayuda de la tabla se calcula la covarianza en x e y. La covarianza, σ_{xy} , de una variable bidimensional (X, Y), es la media del producto de las desviaciones de cada variable en relación con sus respectivas medias.

$$\sigma_{xy} = \frac{\sum(x_i - \bar{x}) \cdot (y_i - \bar{y})}{N} = \frac{\sum x_i y_i f_i}{N} - \bar{x}\bar{y}$$

$$\sigma_{xy} = \frac{34250}{5} - (167{,}6) \cdot (40{,}8) = 11{,}92$$

Observamos que la covarianza es positiva lo que significa que cuando una variable crece, la otra variable también. Las variables tienen una relación directa.

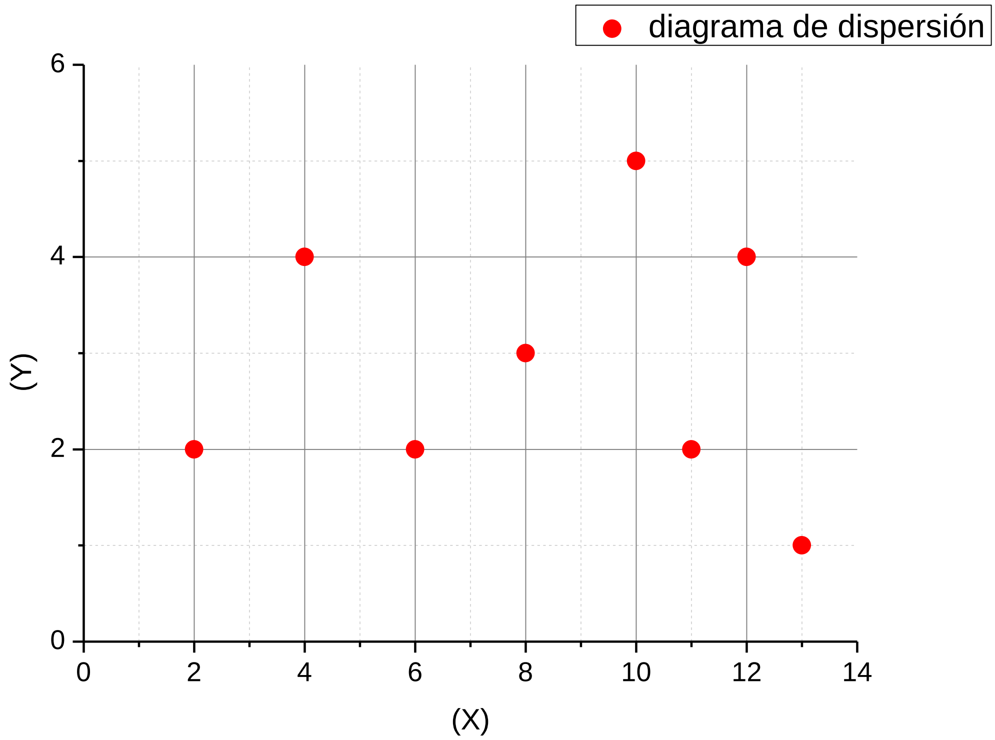
f) El coeficiente de correlación estudia la relación entre las dos variables:

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, con -1 \leq r \leq 1 \ \rightarrow \ r = \frac{11{,}92}{(5{,}004) \cdot (2{,}482)} = \ 0{,}9597 \approx 0{,}96$$

El coeficiente de correlación r es positivo y muy próximo a 1, la correlación entre las variables es fuerte y esto significa que a mayor altura, mayor talla de calzado se usa.

Problema 8. (Oxford Educación – enunciado adaptado del problema 37)

Observa el diagrama de dispersión:



a) Crea la tabla de doble entrada y las tablas de frecuencias marginales.

b) Calcula la varianza de x, la varianza de y, y la covarianza de (x, y).

c) Calcula e interpreta el coeficiente de correlación r.

Solución:

a)

x_i	2	4	6	8	10	11	12	13
y_i	2	4	2	3	5	2	4	1

x_i	2	4	6	8	10	11	12	13
f_i	1	1	1	1	1	1	1	1

y_i	2	4	3	5	1
f_i	3	2	1	1	1

b) Construimos la tabla:

x_i	y_i	f_i	$x_i f_i$	$x_i^2 f_i$	$y_i f_i$	$y_i^2 f_i$	$x_i y_i f_i$
2	2	1	2	4	2	4	4
4	4	1	4	16	4	16	16
6	2	1	6	36	2	4	12
8	3	1	8	64	3	9	24
10	5	1	10	100	5	25	50
11	2	1	11	121	2	4	22
12	4	1	12	144	4	16	48
13	1	1	13	169	1	1	13
66	23	8	66	654	23	79	189

Con ayuda de la tabla se calcula la varianza en x , la varianza en y, y la covarianza de (x, y).

$$\sigma_x^2 = \frac{654}{8} - (8{,}25)^2 = 13{,}6875$$

$$\sigma_y^2 = \frac{79}{8} - (2{,}875)^2 = 1{,}6093$$

$$\sigma_{xy} = \frac{189}{8} - (8{,}25) \cdot (2{,}875) = -0{,}09375 \approx -0{,}094$$

Observamos que la covarianza es negativa lo que significa que cuando una variable crece, la otra variable decrece. Las variables tienen una relación inversa.

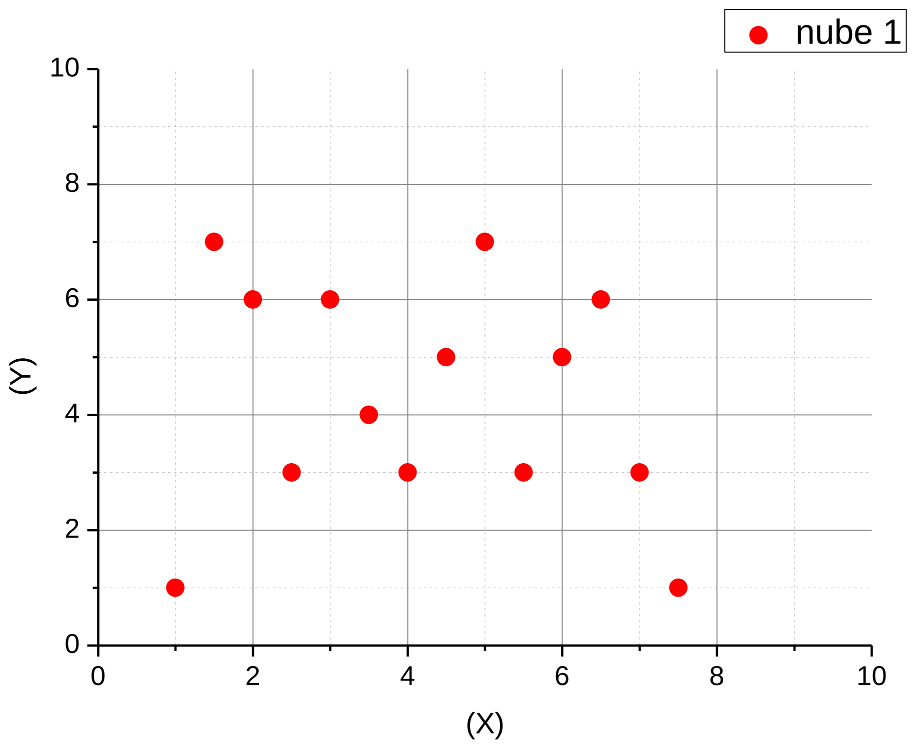
c) El coeficiente de correlación estudia la relación entre las dos variables:

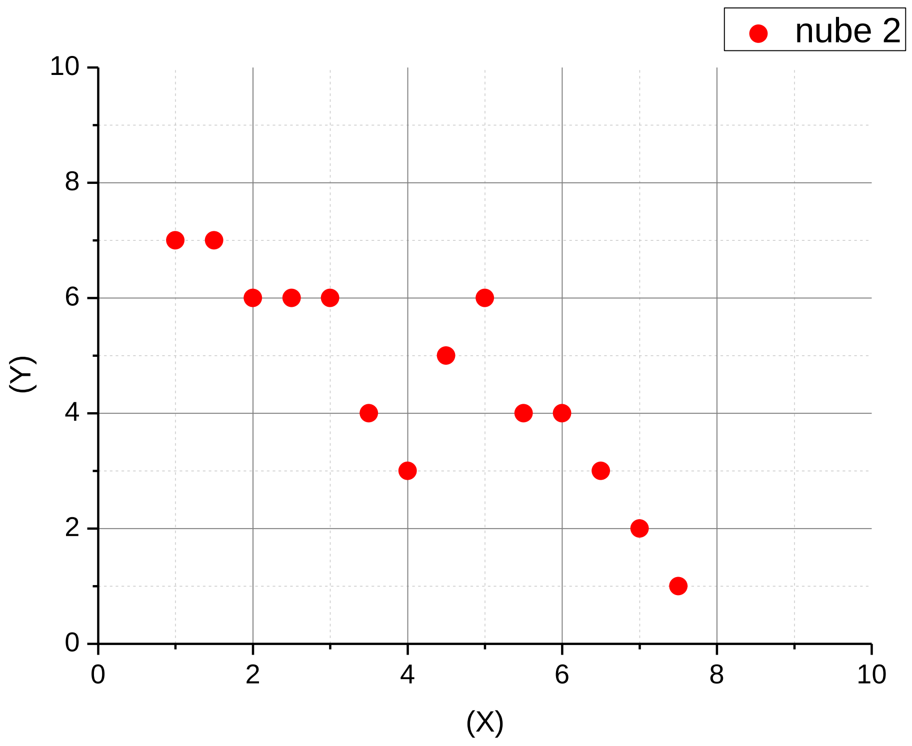
$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, con -1 \leq r \leq 1 \;\rightarrow\; r = \frac{-0{,}094}{(3{,}699) \cdot (1{,}268)} = \; -0{,}02$$

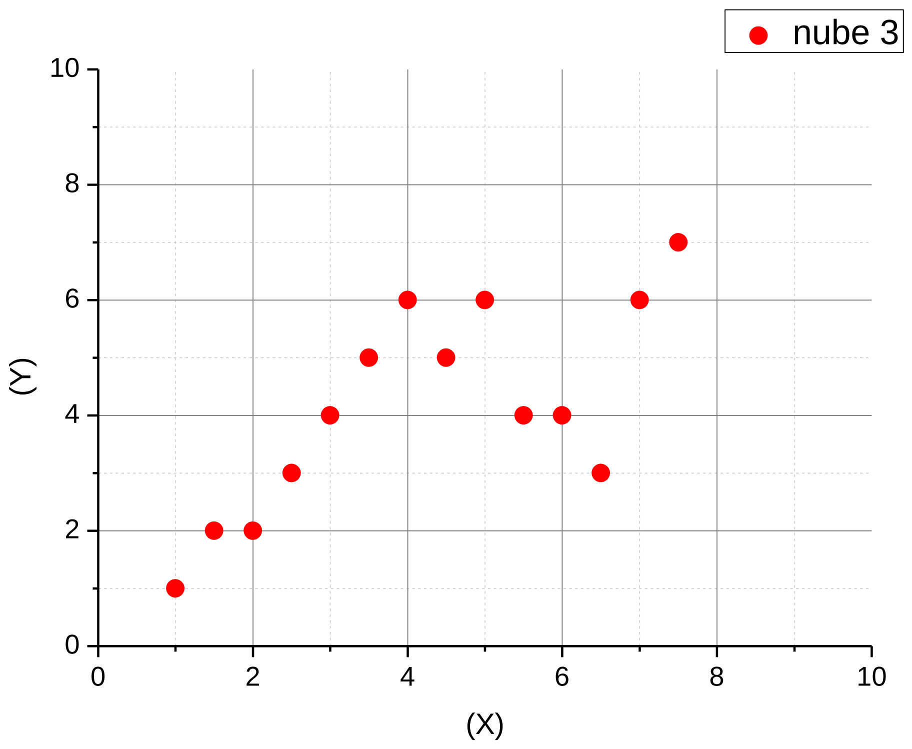
En este caso, aunque el coeficiente de correlación r es negativo, su valor es prácticamente cero lo que implica que no existe correlación entre las variables x e y.

<u>Problema 9.</u> (Savia SM – enunciado del problema 41)

Asocia los coeficientes de correlación r [$r = 0,75$; $r = -0,1$; $r = -0,9$], con su nube de puntos:




nube 2
(Y)
(X)
0
2
4
6
8
10


nube 3
(Y)
(X)
0
2
4
6
8
10

Solución:

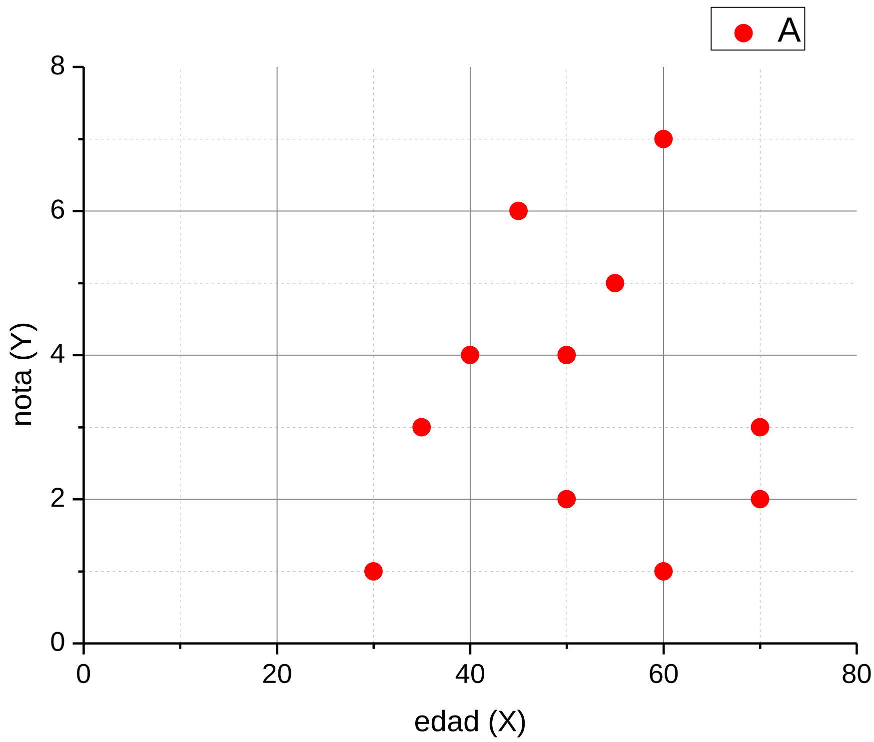
La primera nube de puntos no se ajusta a ninguna recta, por lo que su coeficiente de correlación lineal debe ser $r = -\ 0{,}1$. No existe correlación entre las dos variables.

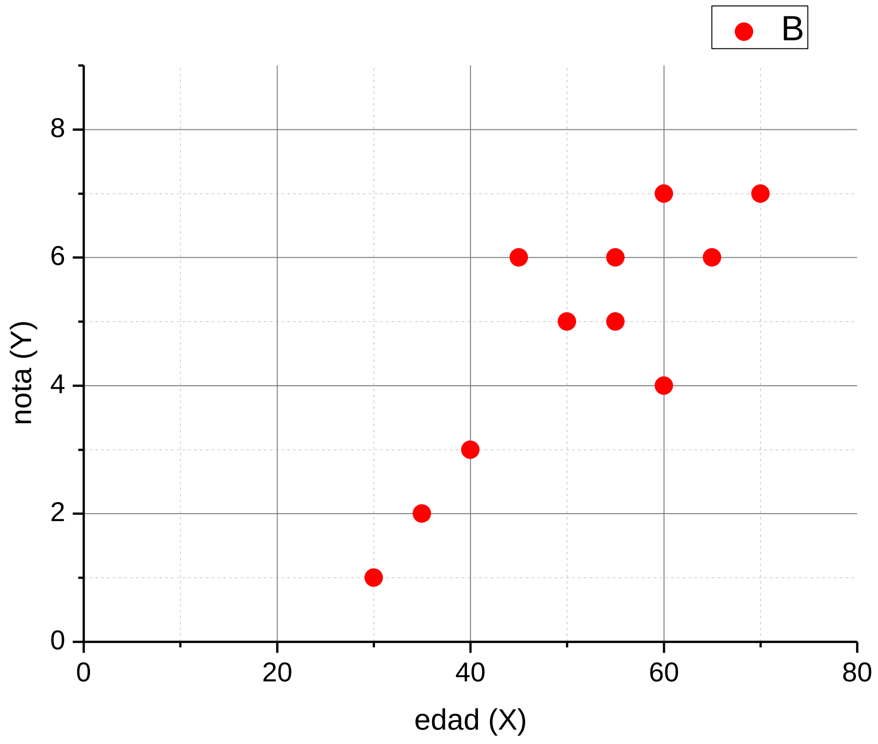
La segunda nube se ajusta a una recta de pendiente negativa, por lo que su coeficiente de correlación lineal es $r = -\ 0{,}9$. Existe una correlación fuerte e inversa entre las dos variables.

La tercera nube se ajusta a una recta con pendiente positiva, por lo que su coeficiente de correlación lineal es $r = 0{,}75$. Existe una correlación moderada y directa entre las dos variables.

Problema 10. (Savia SM – enunciado del problema 45)

¿Cuál de estas dos nubes de puntos representa la edad (X) y las notas obtenidas (Y) en un examen de acceso a la universidad de mayores de 25 años? ¿Por qué?





<u>Solución:</u>

La nube de puntos A no se ajusta a ninguna recta y la nube de puntos B se ajusta a una recta con pendiente positiva. **No existe relación entre la edad (X) y las notas obtenidas (Y) en un examen de acceso a la universidad de mayores de 25 años por lo que la nube de puntos A representa estas dos variables.**

<u>Problema 11.</u> (Paradojas de Martin Gardner)

***Aprende a interpretar el coeficiente de correlación y la relación causal entre variables.* El hecho de que haya correlación fuerte entre dos variables no siempre significa que haya relación causa-efecto entre ellos. Debate sobre las siguientes afirmaciones:**

a) Las estadísticas demuestran que casi todos los accidentes de circulación se producen entre vehículos que ruedan a velocidad moderada. Muy pocos ocurren a mas de 150 km por hora. ¿Significa esto que resulta más seguro conducir a gran velocidad?

b) Un estudio hizo ver que en cierta población europea se produjo simultáneamente un fuerte crecimiento de la población y un notable incremento

del número de cigüeñas. ¿No es esto demostración de que son las cigüeñas quienes traen a los niños al mundo?

c) Un reciente estudio psicopedagógico ha mostrado que los niños de pie grande saben leer mejor que los de pie pequeño. ¿Permitirá el tamaño del pie medir la capacidad de lectura de los niños?

Explicación:

a) No, de ninguna manera. Con frecuencia, las correlaciones estadísticas no reflejan causas y efectos. Casi todo el mundo circula a velocidad moderada, y como es natural, la mayoría de los accidentes se producen a estas velocidades.

b) No; refleja el hecho de que al aumentar el número de edificios las cigüeñas dispusieron de más sitios donde anidar.

c) No, desde luego. El estudio se hizo sobre escolares, que están en crecimiento. Todo cuanto se demostró en él es que los niños mayorcitos, cuyos pies son más grandes, leen mejor que los pequeñines.

Problema 12. (Savia SM – enunciado adaptado del problema 21)

Las edades de los actores que han protagonizado "Romeo y Julieta" en las últimas 9 representaciones de un teatro han sido:

Romeo (X)	**27**	**32**	**39**	**37**	**45**	**38**	**43**	**25**
Julieta (Y)	**23**	**30**	**32**	**40**	**39**	**38**	**32**	**34**

a) Emplea la hoja de cálculo para dibujar la nube de puntos y calcular la recta de regresión.

b) Calcula la media de la edad de Romeo $\bar{x}$ y la media de la edad de Julieta $\bar{y}$.

c) Sitúa en la gráfica el punto $(\bar{x}, \bar{y})$ ¿Qué observas? Sustituye el punto $(\bar{x}, \bar{y})$ en la recta de regresión. De nuevo, ¿qué observas?

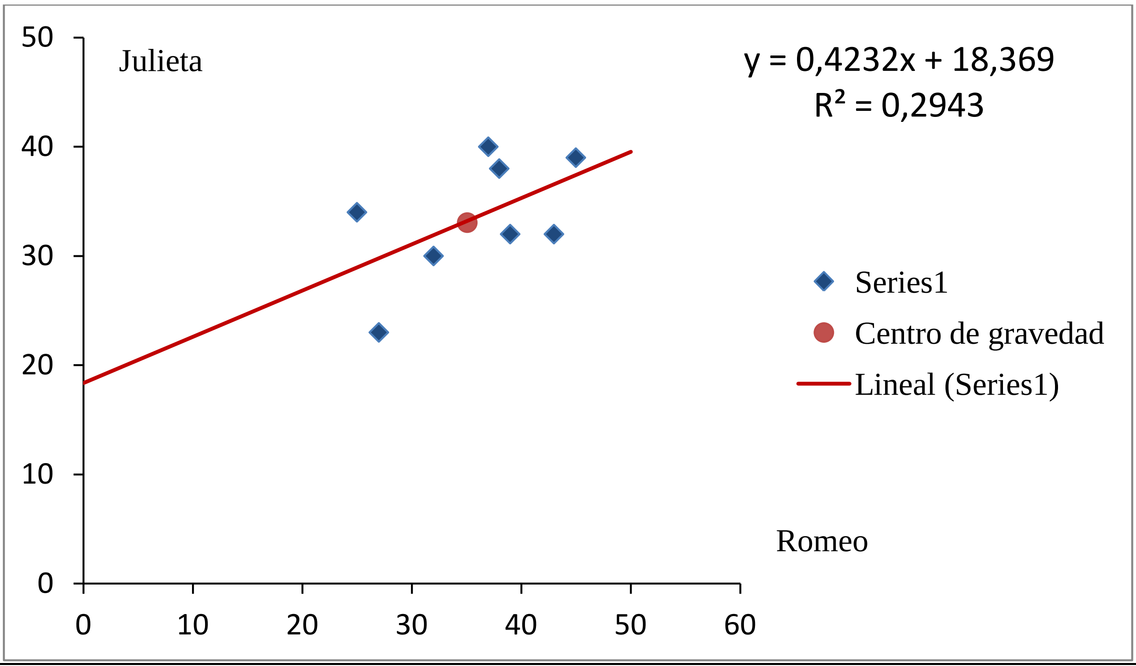
d) En base a los resultados ¿cuál es la edad esperada de Julieta si Romeo tiene 36 años?

Solución:

a) La hoja de cálculo nos ayuda a calcular los parámetros estadísticos, obtener el diagrama de dispersión y calcular la recta de regresión.

	A	B	C	D	E
	Romeo x_i	**Julietay_i**	x_i^2	y_i^2	$x_i y_i$
2	27	23	729	529	621
3	32	30	1024	900	960
4	39	32	1521	1024	1248
5	37	40	1369	1600	1480
6	45	39	2025	1521	1755
7	38	38	1444	1444	1444
8	43	32	1849	1024	1376
9	25	34	625	1156	850
SUMA	**286**	**268**	**10586**	**9198**	**9734**

RECTA DE REGRESIÓN DE Y SOBRE X		
Media (X)=	35,75	"= PROMEDIO (A2: A9)"
Media (Y)=	33,5	"= PROMEDIO (B2: B9)"
Varianza (X)=	45,1875	"= DESVESTP (A2: A9)^2"
Varianza (y)=	27,5	"= DESVESTP (B2: B9)^2"
Desviación típica (X)=	6,72216483	"= DESVESTP (A2: A9)"
Desviación típica (Y)=	5,244044241	"= DESVESTP (B2: B9)"
Media (X.Y)=	1216,75	"= PROMEDIO (E2: E9)"
Covarianza (XY)=	19,125	"= COVAR (A2: A9; B2:B9)"
Coeficiente de correlación=	0,542532725	"= COEF.DE.CORREL(A2: A9; B2:B9)"



b) La media de la edad de Romeo $\bar{x} = 35{,}75$ y la media de la edad de Julieta $\overline{y} = 33{,}5$.

c) Si situamos el punto $(\bar{x}, \overline{y})$ se observa que se encuentra en la recta de regresión. Si se sustituye el punto $(\bar{x}, \overline{y})$ en la ecuación de la recta de regresión observamos que es un punto de la recta.

d) Para hallar la edad esperada de Julieta si Romeo tiene 36 años empleamos la recta de regresión:

$$y - \overline{y} = r \cdot (x - \bar{x})$$

$$Romeo\ 36\ años \rightarrow Julieta:\ y = 0{,}4232 \cdot 36 + 18{,}369 = 33{,}6042\ \approx 34\ años$$

Técnicas o modificaciones de una técnica que se ejercitan

Las técnicas que se ejercitan durante esta sesión son las siguientes:

- establecer una relación entre variables;
- analizar e interpretar una tabla de contingencia;
- representar los valores de las variables (x, y) en un diagrama de puntos bidimensional;
- crear una tabla de contingencia en función del diagrama de dispersión;
- separar una tabla de contingencia, o de doble entrada, en dos tablas simples de frecuencias, una para cada variable;
- representar el centro de gravedad de una distribución bidimensional;
- analizar e interpretar un diagrama de dispersión;
- asociar el coeficiente de correlación a una distribución bidimensional;
- interpretar el coeficiente de correlación y la relación causal entre variables;
- calcular la recta de regresión de y sobre x;
- representar la recta de regresión;
- hacer cálculos predictivos.

Tecnologías y proceso de institucionalización de los distintos aspectos del objeto matemático

Los problemas planteados para estas tres sesiones están diseñados para que el alumno aprenda por descubrimiento los conceptos de la estadística bidimensional. Para ello, deberá poner en funcionamiento sus conocimientos del bloque de funciones, visto anteriormente.

Al ir resolviendo los problemas, será el propio alumno el que institucionalice los siguientes conceptos al descubrir su significado:

- centro de gravedad;
- covarianza;
- coeficiente de correlación;
- recta de regresión lineal.

Metodología para su implementación en el aula

La estadística bidimensional, como he mencionado antes, se introduce por primera vez en secundaria. Por esta razón, el aprendizaje debe ser pausado y progresivo, de modo que el estudiante adquiera las herramientas necesarias para interpretar los parámetros correctamente y estimar la fiabilidad de una previsión mediante un modelo estadístico.

La estadística bidimensional se estudiará durante tres sesiones:

- *en la primera sesión*, al comienzo de clase, se explicarán los objetos matemáticos teóricamente y después se trabajarán los problemas 1-5 por parejas, en el aula ordinaria;
- *en la segunda sesión*, se desarrollarán las técnicas de la primera sesión y se introducirán los nuevos conceptos a medida que se resuelven los problemas 6-10, por parejas también y en el aula ordinaria;
- *la tercera sesión*, que tendrá lugar en el aula de informática, se dividirá en tres partes:
    - *los primeros 10 minutos* estarán destinados a la lectura del problema 11 en grupos cooperativos de cuatro personas, de modo que las conclusiones que extraiga cada grupo se compartirán con el resto de la clase;
    - *los siguientes 30 minutos* para realizar de manera individual el problema 12 con la hoja de cálculo, en este caso el profesor atenderá las dudas de manera individualizada;
    - *los últimos 10 minutos* se emplearán para enunciar los conceptos más importantes de cara al examen y se ofrecerá la hora de tutoría para aclarar las posibles dudas sobre la materia.

El alumno utilizará la calculadora y su cuaderno de clase en todas las sesiones.

6. Cronograma de la secuencia didáctica

La tabla 19 expone el cronograma de la secuencia didáctica, indicando los objetivos y actividades de los campos de problemas (CP), a realizar en cada una de las sesiones.

Tabla 19

Secuencia didáctica y su cronograma

Sesiones	Objetivos didácticos	Actividades didácticas
Sesión 1 50´	Distinguir muestra y población. Reconocer los tipos de variables estadísticas. Organizar datos y frecuencias en tablas.	Problemas 1-7 **(CP 1)** Aula ordinaria. Grupos cooperativos de 4.
Sesión 2 50´	Representar los datos estadísticos en los distintos tipos de gráficos. Interpretar los gráficos.	Problemas 1-7 **(CP 2)** Aula de informática. Por parejas con la hoja de cálculo.
Sesión 3 50´	Calcular e interpretar los parámetros de centralización y posición.	Problemas 1-5 **(CP 3)** Aula ordinaria. Grupos colaborativos de 4.
Sesión 4 50´	Calcular e interpretar los parámetros de dispersión.	Problemas 1-5 **(CP 4)** Aula ordinaria. Grupos colaborativos de 4.
Sesión 5 50´	Construir e interpretar tablas de contingencia. Representar diagramas de dispersión. Deducir la correlación entre dos variables.	Problemas 1-5 **(CP 5)** Aula ordinaria. Por parejas.
Sesión 6 50´	Calcular los parámetros estadísticos bidimensionales. Interpretar los diagramas de dispersión y la correlación entre dos variables.	Problemas 6-10 **(CP 5)** Aula ordinaria. Por parejas.
Sesión 7 50´	Interpretar el coeficiente de correlación y la relación causal entre variables. Calcular e interpretar la recta de regresión.	Problema 11 **(CP 5)**, lectura en grupos cooperativos de 4 (10´). Problema 12 **(CP 5)**, trabajo individual con la hoja de cálculo (30´). Repaso (10´). Aula de informática.

F. Sobre la evaluación

Evaluación del aprendizaje de los alumnos

A continuación, se plantea una prueba escrita de una duración aproximada de una hora, para evaluar el aprendizaje realizado por los alumnos. La prueba consiste en tres problemas que se calificarán según el modelo de tercios del siguiente modo:

- el primer y segundo problema se calificarán sobre dos y tres puntos respectivamente, corresponden a la parte de estadística unidimensional ya estudiada en cursos anteriores;
- el tercer problema se calificará sobre cinco puntos y corresponde a la parte de estadística bidimensional.

Los problemas que se proponen para esta prueba escrita son variaciones de las actividades desarrolladas en las siete sesiones anteriores, de tal modo que **se evalúan todos los estándares de aprendizaje correspondientes al bloque 5 de estadística para este curso.**

Criterios de calificación

Para calificar la prueba escrita se usará un modelo de tercios de modo que:

- por errores en **tareas auxiliares generales** (como el cálculo aritmético o algebraico) no se descontará más de un tercio del valor de la pregunta y se continuará corrigiendo;
- por errores en **tareas auxiliares específicas** (relacionadas con lo estudiado este curso pero que no son el objeto principal de la pregunta) no se descontará más de dos tercios del valor de la pregunta y se continuará corrigiendo;
- por errores en **tareas principales** de la pregunta la penalización será total y se dejará de corregir.

Los objetivos del aprendizaje

Con la realización de la prueba escrita, una vez terminada la secuencia didáctica, se espera que el alumno haya adquirido un buen nivel de conocimiento en los siguientes aspectos estadísticos para dar una respuesta a las preguntas que se plantean:

- identificar los tipos de variable estadística;
- construcción e interpretación de los distintos gráficos estadísticos;
- organizar los datos estadísticos en tablas;
- calcular, interpretar y representar gráficamente los parámetros estadísticos unidimensionales y bidimensionales;
- reconocer la semiótica de los objetos matemáticos en estadística.

La prueba escrita está enfocada a evaluar el sentido estadístico del alumno y el análisis inferencial que hace de las variables. Por este motivo, la prueba se ha diseñado para no perder tiempo en la construcción de las tablas correspondientes a la resolución de cada problema; se proporcionan las tablas y el alumno solo debe elegir la adecuada.

Prueba escrita

Para la realización de la prueba se utilizará calculadora. Se proporcionará a cada alumno tres o cuatro hojas cuadriculadas para responder las preguntas y aparte, el enunciado de los siguientes problemas.

Problema 1. (2 puntos) – (Savia SM – problema de autoevaluación adaptado)

En una encuesta realizada a 500 personas, se ha preguntado el número de periódicos comprados a lo largo de una semana con los siguientes resultados expresados en porcentaje de encuestados:

Nº de periódicos	0	1	2	3	4	5	6	7
Personas (%)	8	15	23	17	12	11	9	5

a) Calcula la mediana y los cuartiles. (1 punto)
b) Halla el recorrido intercuartílico. (0,25 puntos)
c) Representa el diagrama de cajas. ¿Cuántas personas se encuentran por debajo del primer cuartil? (0,75 puntos)

Problema 2. (3 puntos) – (Oxford Educación – problema 53 adaptado)

La tabla muestra el sueldo de los 40 trabajadores de una empresa de transporte de paquetería urgente:

Sueldo (€)	Nº de empleados
[1000, 1400)	4
[1400, 1800)	10
[1800, 2200)	15
[2200, 2600)	11

a) Representa el histograma de frecuencias absolutas. (0,5 puntos)

b) Determina el sueldo medio de los trabajadores, el intervalo mediano y el intervalo modal. (1 punto)

c) Calcula la varianza y la desviación típica. (1 punto)

d) Calcula e interpreta el coeficiente de variación. (0,5 puntos)

Problema 3. (5 puntos) – (Edelvives - problema 38 adaptado)

El Ministerio de Medio Ambiente ha llevado a cabo un estudio sobre el número de incendios forestales que han tenido lugar en una comunidad y las temperaturas medias ambientales registradas durante esos días:

Nº de incendios (X)	2	6	9	5	3	1	6
Tª media en ºC (Y)	30	35	39	34	32	29	36

a) Representa el diagrama de dispersión. (0,5 puntos)

b) Halla el centro de gravedad y sitúalo en el diagrama. (0,5 puntos)

c) Determina e interpreta el coeficiente de correlación lineal. (2 puntos)

d) Calcula la recta de regresión de y sobre x. (1 punto)

e) Si se produjeran 12 incendios, ¿qué presión de temperatura media se podría realizar? (0,5 puntos)

f) ¿Es la previsión fiable? Razona la respuesta. (0,5 puntos)

Nota:

Ten en cuenta las siguientes observaciones, elige la tabla adecuada para resolver cada problema e indica qué tipo de variable es x, y :

$$x_i \rightarrow dato\ x$$
$$y_i \rightarrow dato\ y$$
$$f_i \rightarrow frecuencia\ absoluta$$
$$F_i \rightarrow frecuencia\ absoluta\ acumulada$$
$$\Sigma \rightarrow suma$$

Tabla 1 → Problema…….

	$\boldsymbol{x_i}$	$\boldsymbol{y_i}$	$\boldsymbol{f_i}$	$\boldsymbol{x_i f_i}$	$\boldsymbol{x_i^2 f_i}$	$\boldsymbol{y_i f_i}$	$\boldsymbol{y_i^2 f_i}$	$\boldsymbol{x_i y_i f_i}$
$\Sigma =$								

Tabla 2 → Problema…….

	x_i	f_i	F_i
$\Sigma =$			

Tabla 3 → Problema…….

	x_i	f_i	$x_i f_i$	F_i	$x_i^2 f_i$
$\Sigma =$					

G. Bibliografía

Alcaide, F., Donaire, J. J., Hernández, J., Moreno, M., Pérez, A. & Serrano, E. (2016). *Matemáticas orientadas a las enseñanzas académicas 4 ESO.* Ciudad de México, México: Savia SM.

Aparicio, A. (2018). *Máquina de Galton en Excel.* Recuperado el 25 de Abril de 2020, de https://www.excelavanzado.com/2018/05/maquina-de-galton-en-excel.html

Arnal Bailera, A. & Beltrán Pellicer, P. (2020). Diseño de actividades de aprendizaje de matemáticas. [Apuntes académicos]. ADDUnizar.

Batanero, C., Godino, J. D., Green, D. R., Holmes, P. & Vallecillos, A. (1994). Errores y dificultades en la comprensión de los conceptos estadísticos fundamentales. *Internation Journal of Mathematics Education in Science and Technology, 25*(4), 527-547.

Batanero, C. (2002). Estadística y didáctica de la matemática: Relaciones, problemas y aportaciones. En M. C. Penalva, G. Torregrosa & J. Valls (Coords.), *Aportaciones de la didáctica matemática a diferentes perfiles profesionales* (págs. 95-210). Alicante, España: Universidad de Alicante.

Batanero, C. & Díaz, C. (2004). El Papel de los Proyectos en la Enseñanza y Aprendizaje de la Estadística. En J. Patricio Royo (Ed.), *Aspectos didácticos de las matemáticas* (págs. 125-163). Zaragoza, España: ICE.

Batanero, C. (2004). Los Retos de la Cultura Estadística. *Yupana, 1*(1), 27-37.

Batanero, C. & Godino, J. D. (2005). Perspectivas de la educación estadística como área de investigación. En R. Luengo (Ed.), *Líneas de investigación en Didáctica de las matemáticas* (págs. 203-226). Badajoz, España: Universidad de Extremadura.

Batanero, C., Díaz, C., Contreras, J. M. & Roa, R. (Julio de 2013). El sentido estadístico y su desarrollo. *NÚMEROS, 83*, 7-18.

Batanero, C. & Begué, Nuria. (2017). Significado del muestreo en el currículo de Educación Secundaria Obligatoria. *Segundo Congreso International Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos.*

Cañas Escamilla, J. J. & Galo Sánchez, J. R. (s.f.). *Estadística, Combinatoria y Probabilidad en ESO.* Recuperado el 25 de Abril de 2020, de https://proyectodescartes.org/iCartesiLibri/materiales_didacticos/IntroduccionEstadisticaProbabilidad/index.html

Colera, J., de Guzman, M. & Salvador, A. (1987). *MATEMATICAS. Bachillerato 1*. Madrid, España: Ediciones Anaya, S.A.

de Lucas Benedicto, M. & Rey Fedriani, J. M. (2016). *Matemáticas orientadas a las enseñanzas académicas 4 ESO, volumen 3*. Madrid, España: Oxford University Press España, S. A.

Estadística para todos. (s.f.). *Historia de la Estadística*. Recuperado el 25 de Abril de 2020, de https://www.estadisticaparatodos.es/historia/histo_esta.html

Estadística para todos. (s.f.). *Hoja de Cálculo Excel/Calc*. Recuperado el 25 de Abril de 2020, de https://www.estadisticaparatodos.es/software/excel.html

Estadística para todos. (s.f.). *Otros software estadísticos*. Recuperado el 25 de Abril de 2020, de https://www.estadisticaparatodos.es/software/software_otros.html

Fernández Bravo, J. A. (2014). *Cuaderno 3 de Matemáticas 3º Primaria Superpixépolis*. Zaragoza, España: Edelvives.

Fernández Palop, M. P. & Caballero García, P. A. (2017). El libro de texto como objeto de estudio y recurso didáctico para el aprendizaje: fortalezas y debilidades. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 20*(1), 201-217.

Gal, I. (2002). Adult´s statistical literacy. Meanings, components, responsabilities. *International Statistical Review, 70*(1), 1-25.

Gardner, M. (1983). *¡Ajá! PARADOJAS. Paradojas que hacen pensar*. Barcelona, España: LABOR, S.A.

Holmes, P. (2002). Some lessons to be learnt from curriculum developments in statistics. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*. Ciudad del Cabo: IASE. CD ROM.

Hurtado, J. (2013). *BREVE HISTORIA DE LA ESTADÍSTICA*. Universidad Nacional de Colombia, Medellin.

Instituto Nacional de Estadística [ES]. (s.f.). *Historia de la Estadística*. Recuperado el 1 de Abril de 2020, de https://www.ine.es/explica/docs/historia_estadistica.pdf

Ley Orgánica 2/2006, de 3 de mayo, de Educación. Boletín Oficial del Estado. Madrid, 4 de mayo de 2006, núm. 106, pp. 17158-17207.

Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa. Boletín Oficial del Estado. Madrid, 10 de diciembre de 2013, núm. 295, pp. 97858-97921.

Martínez Juste, S. & Muñoz Escolano, J. M. (2019). Diseño curricular e instruccional de matemáticas. [Apuntes académicos]. ADDUnizar.

Martínez Juste, S. & Muñoz Escolano, J. M. (2020). Innovación e investigación educativa en matemáticas. [Apuntes académicos]. ADDUnizar.

Mejía Sánchez-Bermejo, D., Ocaña Fernández, J. M. & Romero Torralba, R. (2016). *ESO 4 matemáticas académicas, volumen 3.* Zaragoza, España: Edelvives.

Occelli, M. & Valeiras, N. (2013). Los libros de texto de ciencias como objeto de investigación: una revisión bibliográfica. *Enseñanza de las Ciencias, 31,* 133-152.

Olave, P. (2019). Big data: La fiebre del siglo XXI. Reinventándose la estadística para los nuevos datos. *conCIENCIAS.digital*(24), 56-67.

Orden de 16 de junio de 2014, de la Consejera de Educación, Universidad, Cultura y Deporte, por la que se aprueba el currículo de la Educación Primaria y se autoriza su aplicación en los centros docentes de la Comunidad Autónoma de Aragón. Boletín Oficial de Aragón. Aragón, 20 de junio de 2014, núm. 119, pp. 19288-20246.

Orden ECD/489/2016, de 26 de mayo, por la que se aprueba el currículo de la Educación Secundaria Obligatoria y se autoriza su aplicación en los centros docentes de la Comunidad Autónoma de Aragón. Boletín Oficial de Aragón. Aragón, 2 de junio de 2016, núm. 105, pp. 12640-13458.

Orden ECD/850/2016, de 29 de julio, por la que se modifica la Orden de 16 de junio de 2014, de la Consejera de Educación, Universidad, Cultura y Deporte, por la que se aprueba el currículo de la Educación Primaria y se autoriza su aplicación en los centros docentes de la Comunidad Autónoma de Aragón. Boletín Oficial de Aragón. Aragón, 12 de agosto de 2016, núm. 156, pp. 20713-20884.

Real Decreto 126/2014, de 28 de febrero, por el que se establece el currículo básico de la Educación Primaria. Boletín Oficial del Estado. Madrid, 1 de marzo de 2014, núm. 52, pp. 19349-19420.

Redal, E. J. (2011). *Matemáticas 4 ESO, opción B, volumen3.* Madrid, España: Santillana Educación, S.L.